

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Martin Šiškovič

**Analiza intonacije petja z algoritmi za
detekcijo višine tona**

DIPLOMSKO DELO

VISOKOŠOLSKI STROKOVNI ŠTUDIJSKI PROGRAM PRVE
STOPNJE RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: viš. pred. dr. Robert Rozman

Ljubljana, 2017

Rezultati diplomskega dela so intelektualna lastnina avtorja in Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

Besedilo je oblikovano z urejevalnikom besedil \LaTeX .

Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Tematika naloge:

V okviru diplomske naloge poiščite vzroke za spreminjanje intonacije pri petju. Petje analizirajte z algoritmi za določanje višine tona. Najprej primerjajte obstoječe metode in izberite najbolj primerne. Zasnujte in implementirajte sistem in z njim po potrebi obdelajte tudi realne posnetke, ki jih sami posnamete. Rezultate primerno grafično prikažite in analizirajte. Na podlagi rezultatov podajte končne ugotovitve in ovrednotite njihovo skladnost z obstoječo literaturo na tem področju.

Hvala družini, ki mi je v času študija stala ob strani: Angeli, Frankotu, Igorju, Marti, Nini, Pii in Radotu. Za pomoč pri diplomski nalogi se zahvaljujem mentorju Robertu Rozmanu. Tini Gustinčič gre zahvala za slovnični pregled.

Za sodelovanje pri praktičnem delu se zahvaljujem članom Akademskega pevskega zbora Univerze na Primorskem. Ambrož Čopi, Ana Štrajhar, Andreja Tomažič Hrvatini, Andreja Trček, Andrej Leban, Anže Bizjak, Davorin Lovrečič, Fabiana Geri, Jan Veljak, Jure Kristan, Karin Kres, Kevin King, Klara Lukan, Klavdija Lukač, Luka Vetrih, Luka Žorž, Martin Marchesich, Patricija Gornjec, Petra Rančov, Sandi Štor, Sanja Đukić, Urška Terbižan in Zala Gruden, hvala!

Svoji družini.

Kazalo

Povzetek

Abstract

1	Uvod	1
2	Metode	5
2.1	Merjenje frekvence signalov	5
2.1.1	Frekvenčna domena	6
2.1.2	Časovna domena	6
2.1.3	Avtokorelacija	8
2.1.4	Središčno izrezovanje	13
2.1.5	Metoda pYIN	14
2.2	Intonacija	15
2.2.1	Naravna uglasitev	17
2.2.2	Temperirana uglasitev	18
2.2.3	Problem točne intonacije	23
3	Uporabljena programska orodja	25
3.1	Programski jezik Python	25
3.2	Paketi in knjižnice	26
3.2.1	Audiolab	26
3.2.2	NumPy	26
3.2.3	SciPy	27

3.2.4	Matplotlib	27
3.3	Audacity	28
3.4	Sonic Visualiser	28
4	Implementacija sistema za analizo posnetkov	29
4.1	Določanje osnovne frekvence	30
4.2	Notni zapis MIDI	35
4.3	Prikaz rezultatov	36
5	Rezultati analize realnih posnetkov	37
5.1	Iskanje napak pri intervalih	37
5.2	Iskanje napak pri razlikah med uglasitvami	42
6	Sklepne ugotovitve	45
6.1	Pregled praktičnega dela in rezultatov	45
6.2	Analiza rezultatov in izboljšave	47
6.3	Praktičen pomen ugotovitev	48
	Literatura	48

Seznam uporabljenih kratic

kratica	angleško	slovensko
F_t	Pitch Frequency	Frekvenca tona
F₀	Fundamental Frequency	Osnovna frekvenca
F_s	Sampling Frequency	Frekvenca vzorčenja
Hz	Hertz	Herc
MIDI	Musical Instrument Digital Interface	Digitalni vmesnik za glas- bene instrumente

Povzetek

Naslov: Analiza intonacije petja z algoritmi za detekcijo višine tona

Splošno znan pojav pri pevcih je spreminjanje intonacije med petjem. Z njim se srečujejo vsi, ki se ukvarjajo s petjem brez instrumentalne spremljave. Kljub njegovi pogostosti in nezaželenosti ni znan enoten razlog za prisotnost pojava. Običajno se zmotno trdi, da je spreminjanje intonacije odvisno od kvalitete pevcev. Cilj diplomske naloge je poiskati vzroke za spremembe v intonaciji. Narejen je bil pregled znanstvenih objav, ki so povezane s tematiko ter analiza sprememb v intonaciji na realnih posnetkih. Posnetke smo sami naredili v okviru diplomske naloge. Ti so bili analizirani z algoritmi za detekcijo višine tona. Uporabili smo obstoječi vtičnik algoritma pYIN in implementirano lastno različico avtokorelacije. Rezultati analize so dodatno potrdili ugotovitve nekaterih objav, ki pravijo, da se intonacija v skladbi spreminja zaradi razlik med naravno in temperirano uglasitvijo. Spremembe se pojavijo, ko je med dvema tonoma razlika med uglasitvama negativna. Več kot je takih primerov v skladbi, večja je lahko sprememba v intonaciji.

Ključne besede: algoritmi za detekcijo višine tona, analiza petja, avtokorelacija, intonacija, sprememba intonacije.

Abstract

Title: Singing Voice Intonation Analysis using Pitch Detection Algorithms

Pitch drift is a commonly known phenomenon among singers. It occurs in most a cappella groups. Reasons for the phenomenon are not known, despite that, it is very common and unwanted. It is usually mistakenly argued that pitch drift depends on the singers' skills. The objective of this thesis is to find the causes of pitch drift. We reviewed scientific publications related to the theme and analyzed the changes in intonation on audio tracks, which we recorded in the context of this thesis. Tracks were analyzed with algorithms for pitch detection. We used an existing plug-in based on pYIN algorithm and our own modified implementation of an algorithm based on autocorrelation. The analysis results have further confirmed the findings of some publications, saying that the intonation in the songs changes due to differences between the just intonation and the equal temperament. Changes occur when difference in tuning between two notes is negative. More such cases are there in the song, greater change in the intonation can occur.

Keywords: pitch detection algorithms, singing analysis, autocorrelation, intonation, pitch drift.

Poglavje 1

Uvod

Glasba se izraža s petjem ali z glasbenimi instrumenti. Ti se med seboj po sestavi lahko zelo razlikujejo, na drugi strani pa je vokalni trakt¹ pri vsakem človeku bolj ali manj enak. Ljudje s petjem ustvarjajo glasbo po celem svetu. Vokalna glasba je prisotna že od prazgodovinskih časov in je bila mogoče prisotna še pred nastankom prvih govornih jezikov [1]. Pri zborovskem petju je intonacija glavna prioriteta. V sestavih brez glasbene spremljave se pogosto opazi spreminjanje intonacije. Največkrat je problem predvsem pri spreminjanju intonacije navzdol² [1].

Motivacija za delo se je pojavila na podlagi izkušenj pri petju v zboru. Pogosto smo se srečevali s spreminjanjem intonacije, nikoli pa nismo razumeli razlogov za ta pojav. V diplomski nalogi želimo raziskati ozadje in poiskati vzroke za pojav, ki je v zborovskem petju pogost in hkrati izredno nezaželen. Naredili smo pregled nad literaturo, ki opisuje problem spreminjanja intonacije. Za potrebe naloge smo napisali tri programe. V prvem je implementiran algoritem za določanje višine tona, druga dva pa analizirata dobljene rezultate in jih grafično prikažeta. S programi smo analizirali posnetke pevcev Akademskega pevskega zbora Univerze na Primorskem, ki smo jih sami naredili v okviru tega diplomskega dela.

¹Vokalni trakt sestavljajo: žrelo, požiralnik, grlo, sapnik, nosna in ustna votlina.

²Pojav, kjer je zadnji ton v skladbi nižje uglasen od prvega.

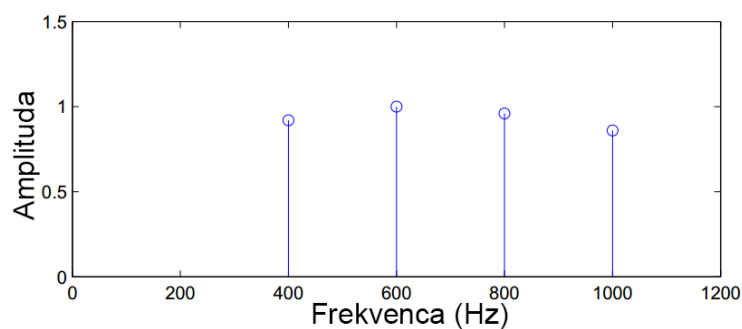
Intonacijo je mogoče razumeti kot točnost note pri petju ali igranju na instrument. Teorija nakazuje obstoj referenčnega tona, od katerega se meri napako (relativno razliko v poltonih). Dandanes obstaja univerzalen način za izbiro referenčnega tona. Po standardu je srednji ton A (A_4) uglasen na 440 Hz. Na to frekvenco so največkrat uglasene glasbene vilice. To pa ni od vselej veljalo za standard. Orgle so v določenih obdobjih imele uglasen ton A_4 od frekvence 374 do 567 Hz. Glasbene vilice iz zgodnjega 18. stoletja so bile uglasene na 422,5 Hz. To je veljalo kot standard za dve stoletji, v času ko so bila napisana dela Mozarta, Bacha in Beethovna. Ta dela se danes izvajajo skoraj polovico tona višje. Višina tona se je z leti spreminjala zaradi želje glasbenikov po svetlejšem zvoku [2].

Višino tona si lahko razlagamo kot dožemanje visokega ali nizkega zvoka. Ton je povezan s frekvencami v zvoku. Prisotnost višjih frekvenc zaznamo kot višji ton. V literaturi se pojavljata pojma frekvenca tona in osnovna frekvenca. Frekvenca tona (F_t) je definirana kot frekvenca čistega sinusnega signala, katerega višina je enaka višini tona. Osnovna frekvenca (F_0) je definirana kot inverz dolžine periode tona, kjer periodo predstavlja najmanjša ponavljajoča se enota signala. V harmoničnem signalu je to najnižja frekvenca v harmonični vrsti. Oba pojma pogosto sovpadata in predpostavlja se, da sta enaka. Med komponentami zvoka pa ni nujno vedno prisotna osnovna frekvenca. V tem primeru jo lahko določimo iz njenih večkratnikov - ostalih frekvenčnih komponent.

Slika 1.1 prikazuje tak primer. Osnovno frekvenco valovne oblike določa razdalja med harmoniki³. Če jo izračunamo, ugotovimo, da znaša 200 Hz. Osnovna frekvenca signala je torej 200 Hz, čeprav je amplituda frekvence 200 Hz enaka 0. Ob naraščanju frekvence se manjša število harmonikov, s katerimi lahko določimo višino tona. Nad 2500 Hz se za določanje tona uporablja samo osnovna frekvenca.

Prvi poskusi zaznavanja frekvence so bili izvedeni z resonatorji. Helmholtz je okoli leta 1860 odkril način za detekcijo frekvence. Resonatorji so

³Harmoniki ali alikvotni toni so frekvence, ki so celi večkratniki osnovne frekvence.



Slika 1.1: Primer frekvenčne sestave zvoka brez osnovne frekvence [2].

bili običajno izdelani iz stekla ali medenine, sferične oblike z ozko odprtino. Taka naprava, ki jo prikazuje slika 1.2, ima svojo resonančno frekvenco.



Slika 1.2: Helmholtzov resonator [2].

Ko se ob napravi oglasi zvok, ki vsebuje to frekvenco, naprava resonira. Kot rezultat proizvaja zvok, ki poudari to frekvenco. To je mogoče slišati, če se resonator prisloni k ušesu. Tako lahko tudi amater presodi, če je v zvoku

prisotna določena frekvenca. Resonančna frekvenca resonatorja je določena z naslednjim izrazom:

$$f = \frac{v}{2\pi} \sqrt{\frac{a}{Vl}}, \quad (1.1)$$

kjer a označuje prerez vratu, l je dolžina vratu, V je notranja prostornina, v pa hitrost zvoka. Izdelati je mogoče mnogo resonatorjev različnih velikosti, za različne frekvence. Helmholtz je s temi resonatorji dokazal, da so kompleksni toni sestavljeni iz niza različnih frekvenc. Frekvenčne komponente se imenujejo alikvotni toni⁴. To so toni, ki so skoraj neslišni in skupaj z osnovno frekvenco tvorijo kompleksni ton. Od njih je odvisna barva zvoka. Nekateri instrumenti (npr. strunski) ustvarijo harmonične alikvotne tone, nasprotno pa recimo zvonovi ustvarijo alikvotne tone, ki niso harmonični. Človeški glas pri petju ustvari harmonične alikvotne tone.

V naslednjem poglavju so opisane metode za merjenje frekvence signalov. Nekatere izmed njih so tudi podrobno predstavljene. Na koncu poglavja obrazložimo še intonacijo in dva različna načina uglasitve. Tretje poglavje opisuje programska orodja, ki smo jih uporabili pri diplomski nalogi. V četrtem poglavju opišemo implementacijo našega sistema za analizo posnetkov.

Sledi poglavje s praktičnim delom, ki je razdeljeno na dva dela. V prvem delu se nahaja analiza posnetkov posameznih pevcev, ki so samostojno odpeli melodijo v izoliranem prostoru. Izbrali smo dve melodiji. Ena vsebuje velike razmike med notami, pri drugi pa se določeni razmiki ponavljajo čez celo skladbo. V analizi posameznih pevcev preverjamo, če je lahko za spremembo v intonaciji kriva napaka pri določenih razmikih med notami. V drugem delu so analizirani posnetki štirih pevcev, ki sočasno odpojejo pevsko vajo. Pri tej vaji naj bi bili premiki v intonaciji izrazitejši. Ta del je namenjen iskanju povezave med spreminjanjem intonacije in razlikami med različnima uglasitvama.

V zadnjem poglavju podamo še sklepne ugotovitve.

⁴Imenujemo jih tudi: sozveneči toni, sotoni, parcialni toni ali višje harmonski toni.

Poglavje 2

Metode

Večino praktičnega dela diplomske naloge predstavlja analiza posnetkov. Te smo primerjali z notnim zapisom. Za primerjavo smo morali posnetke pretvoriti v drugačno obliko. Vsak ton v posnetku je definiran s točno določeno frekvenco F_0 . Iz teh frekvenc smo izračunali MIDI vrednosti¹. Te smo nato primerjali z MIDI vrednostmi notnega zapisa.

V tem poglavju je opisano delovanje nekaterih metod za izračun osnovne frekvence. Dve izmed njih nato podrobneje predstavimo. Na koncu se dotaknemo še različnih uglasitev in problema petja v točni intonaciji.

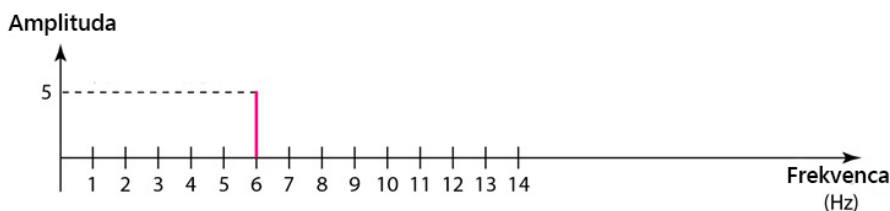
2.1 Merjenje frekvence signalov

Signal, ki ga sproducira instrument ali glas ni nikoli periodičen v ožjem matematičnem smislu, je pa zelo blizu približni oziroma t.i. kvazi periodičnosti. Zvočni signal si je mogoče razlagati v splošnem kot signal, katerega amplituda in frekvenca se v času spreminjata. Pogosto je tak signal digitalno posnet govor ali petje. Pri analizi tovrstnih posnetkov se srečamo z metodami, ki delujejo v časovni ali frekvenčni domeni [3], ki sta opisani v nadaljevanju. Sledijo še opisi izbranih metod za merjenje frekvence signalov.

¹MIDI vrednosti so številčne oznake not. Podrobna razlaga se nahaja v podpoglavju 4.2.

2.1.1 Frekvenčna domena

V to domeno prištevamo metode s Fourierjevo transformacijo, ki preslikajo signal iz časovnega v frekvenčni prostor [2]. Pri tej preslikavi lahko v dobljenem spektru vidimo prisotne frekvence v zvoku in njihove amplitude (slika 2.1). Med vrhovi je izbran najboljši kandidat, ki je predstavnik osnovne frekvence. Te metode so učinkovite tudi ob zelo nizkem razmerju signal/šum in ob šumu iz ozadja [3].



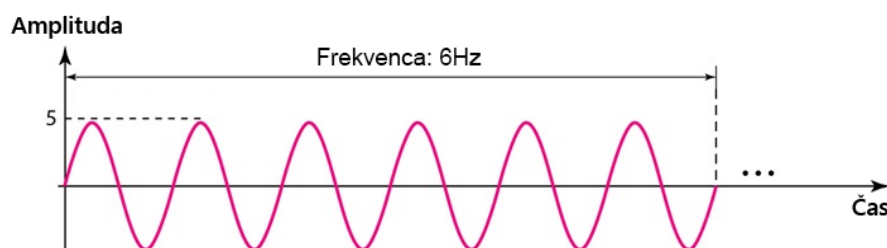
Slika 2.1: Sinusni signal 6 Hz v frekvenčni domeni [4].

2.1.2 Časovna domena

Metode v časovni domeni obdelujejo neposredno zapise signalov, ki jih dobimo z vzorčenjem. To so običajno kar zaporedja zajetih vzorcev signala. V njih enakomerno razporejeni vzorci predstavljajo potek amplitude zvoka v času [2]. Metode frekvenco izračunajo iz periode signala z naslednjim izrazom:

$$\text{frekvenca} = \frac{1}{\text{perioda}}, \quad (2.1)$$

kjer je perioda izražena v času. Slika 2.2 prikazuje potek sinusnega signala v časovni domeni. Puščica označuje čas 1 sekunde na vodoravni osi, v kateri se zgodi 6 period signala.



Slika 2.2: Sinusni signal 6 Hz v časovni domeni [4].

Najpogosteje je uporabljena frekvenca vzorčenja F_S 44100 vzorcev na sekundo. Metode se dobro obnesejo pri zelo periodičnih signalih. Zelo občutljive so na šum ter nagnjene k napakam, kjer je kot rezultat namesto osnovne frekvence izbran njen prvi alikvotni ton [3]. Najbolj znane metode v časovni domeni so [5]:

- "zero-crossing",
- avtokorelacija,
- diferenčne enačbe,
- izpeljane metode, ki temeljijo na avtokorelaciji ali diferenčni enačbi:
 - YIN,
 - pYIN,
 - RAPT,
 - PRAAT,
 - AMDF.

V diplomski nalogi smo v prvi polovici praktičnega dela analizirali posnetke posameznih pevcev, ki so samostojno odpeli melodijo. Pri teh posnetkih je za meritev frekvenčne sestave uporabljen pYIN algoritem, v obliki obstoječega Vamp vtičnika [6] za orodje Audacity. V drugi polovici so analizirani posnetki štirih pevcev, ki sočasno odpojejo pevsko vajo. Ti posnetki so

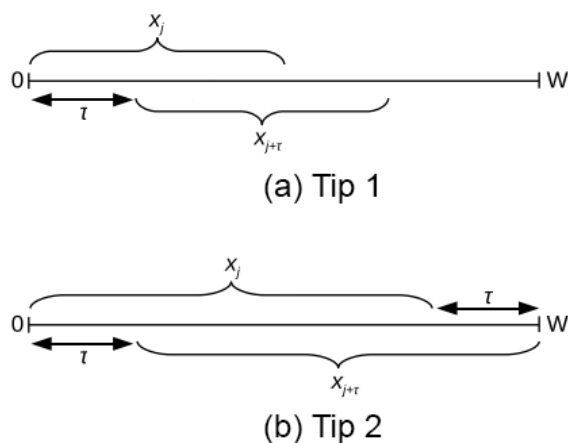
analizirani z metodo avtokorelacije, ki smo jo implementirali v programskem jeziku Python. Implementacija je opisana v podpoglavju 4.1.

2.1.3 Avtokorelacija

Avtokorelacija je ena izmed enostavnejših metod za detekcijo višine tona v časovni domeni [7]. Ta operacija pravzaprav pomeni izračun križne korelacije signala s svojo lastno kopijo pri različnih zamikih [2]. Vsak vzorec se pomnoži z zamaknjeno kopijo in se sešteje v končni rezultat avtokorelacije pri določenem odmiku. Avtokorelacijo $r(\tau)$ zapišemo z izrazom:

$$r(\tau) = \sum_{j=-\infty}^{\infty} x_j x_{j+\tau}. \quad (2.2)$$

Če je signal periodičen s periodo p , velja $x_j = x_{j+p}$. Avtokorelacija bo imela lokalne maksimalne vrednosti pri odmikih, ki so večkratniki števila p . Tam se funkcija najboljše ujema sama s seboj (slika 2.4).



Slika 2.3: Definiciji avtokorelacije tipa 1 in 2 [2].

Ko se avtokorelacijo izvaja na posnetkih, se uporablja kratke odseke, saj se frekvenčna vsebina signalov spreminja. Ti odseki signala se imenujejo

okvirji. Njihova dolžina je običajno dovolj kratka, da se frekvenčna sestava znotraj njih ne spreminja preveč. Lahko se zgodi, da so zaradi prevelike velikosti okvirja pri visokih frekvencah približki manj natančni. Avtokorelacija bo v tem primeru pokazala povprečno sliko signala v okvirju. Vendar pa mora biti okvir dovolj dolg, da dobimo smiselni rezultat korelacije. Vsebovati mora vsaj dve periodi signala.

Avtokorelacijo okvirja je mogoče definirati na dva načina (slika 2.3). Imenovali ju bomo tip 1 in tip 2. Tip 1 podaja izraz 2.3, kjer je avtokorelacija $r(\tau)$ zakasnjena za τ in izračunana za okvir dolžine W . Vsota vedno vsebuje $W/2$ elementov in je zakasnjena za največ $\tau=W/2$.

$$r(\tau) = \sum_{j=0}^{W/2-1} x_j x_{j+\tau}, \quad 0 \leq \tau \leq W/2 \quad (2.3)$$

Avtokorelacija tipa 2 $r'(\tau)$ je definirana z izrazom:

$$r'(\tau) = \sum_{j=0}^{W/2-1-\tau} x_j x_{j+\tau}, \quad 0 \leq \tau < W. \quad (2.4)$$

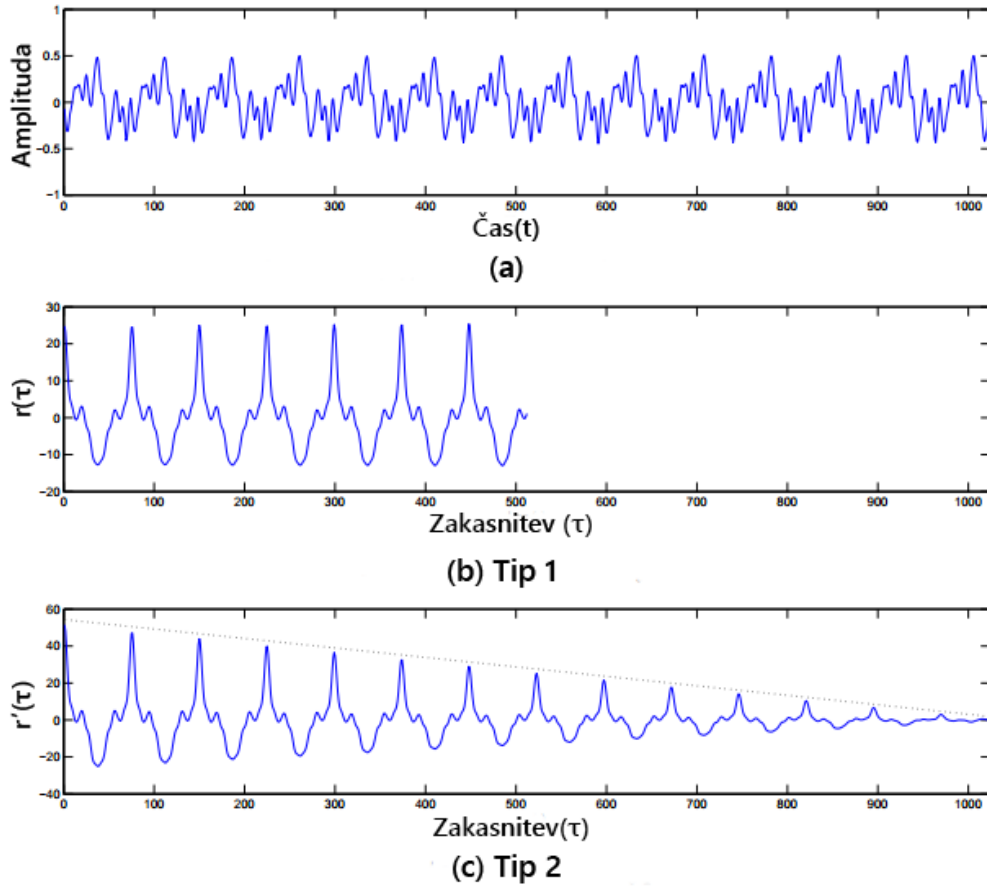
Pri avtokorelaciji seštevamo pomnožene vrednosti med vzorci nezakasnjene in zakasnjene signala. Z večanjem zakasnitve τ se število med seboj pomnoženih elementov zmanjšuje. To se na sliki 2.4(c) opazi kot linearno zmanjševanje vrednosti proti ničli v točki $\tau = W$. Linearno zmanjševanje vrednosti je mogoče odpraviti z nepristransko obliko², ki jo zapišemo z izrazom:

$$r(\tau) = \frac{W}{W-\tau} \sum_{j=0}^{W/2-1-\tau} x_j x_{j+\tau}, \quad 0 \leq \tau < W. \quad (2.5)$$

Slika 2.4 prikazuje izhodne podatke obeh tipov avtokorelacije nad okvirjem posnetka violine. Večje nestabilnosti se lahko začnejo pojavljati, ko se

²Angl. "Unbiased form".

τ približuje W , zaradi tega so uporabljene izhodne vrednosti samo do $W/2$. Maksimum funkcije $r'(\tau)$ je dosežen v $\tau = 0$, kar je sorazmerno z energijo podatkov v okvirju. To maksimalno vrednost je mogoče uporabiti pri normalizaciji. Vse vrednosti delimo z $r'(0)$ in dobimo le vrednosti med -1 in 1.

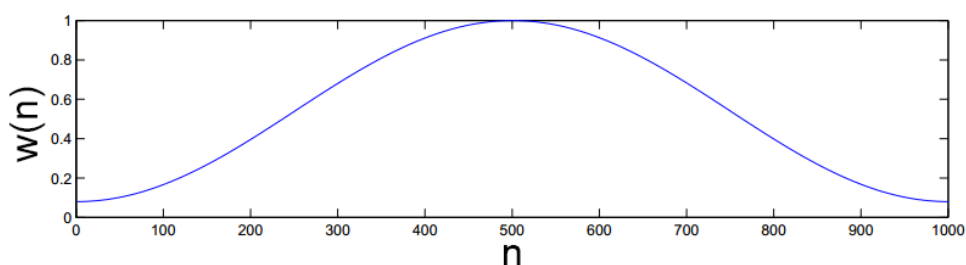


Slika 2.4: Del posnetka violine (a), oba tipa avtokorelacije (b),(c) [2].

Pred avtokorelacijo se podatke razreže v posamične okvirje. Nenaden začetek ali konec avtokorelacije pri robu okvirja povzroči robni učinek. Učinek je manjši, če so vrednosti višje na sredini okvirja. Tako je premikanje okvirja bolj gladko. Za zmanjševanje robnega učinka se pogosto uporablja množenje vzorcev v okvirju z vzorci Hammingovega okna, ki se jih izračuna z izrazom:

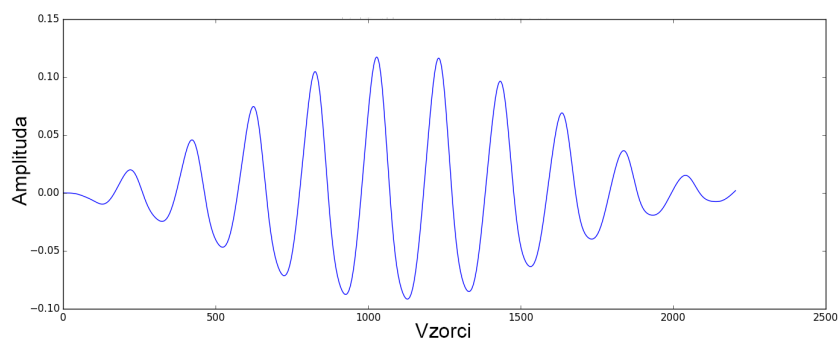
$$w_n = 0.53836 - 0.46164 \cos\left(\frac{2\pi n}{W-1}\right), 0 \leq n < W. \quad (2.6)$$

Funkcija ima ničelne vrednosti izven okvirja.



Slika 2.5: Hammingovo okno dolžine 1000 [2].

Slika 2.5 prikazuje primer Hammingovega okna dolžine $W = 1000$. Okvir se preoblikuje z množenjem po elementih s Hammingovim oknom. Oba morata biti iste dolžine.



Slika 2.6: Okvir pomnožen s Hammingovim oknom.

Tako preoblikovan okvir samoglasnika "a" prikazuje slika 2.6. Dobljene vrednosti se pri robovih nahajajo bližje ničli. Avtokorelacija okvirja preoblikovanega s koeficienti Hammingovega okna je definirana kot:

$$r'_{Win}(\tau) = \sum_{j=0}^{W-1-\tau} w_j x_j w_{j+\tau} x_{j+\tau}. \quad (2.7)$$

Izraz je uporabnejši v nepristranski obliki³. Nepristranska avtokorelacija preoblikovana s Hammingovim oknom je definirana z izrazom:

$$r'_{WinUnbiased}(\tau) = \frac{W}{(W-\tau)} \sum_{j=0}^{W-1-\tau} w_j x_j w_{j+\tau} x_{j+\tau}. \quad (2.8)$$

Tudi v tem primeru je mogoča normalizacija z deljenjem rezultatov z vrednostjo na ničelnem mestu.

Potem ko je bila avtokorelacija izvedena, v njenih rezultatih poiščemo lokacije maksimumov. Razdalja med sosednjima maksimumoma predstavlja oceno osnovne periode. Iz nje lahko izračunamo osnovno frekvenco z naslednjim izrazom:

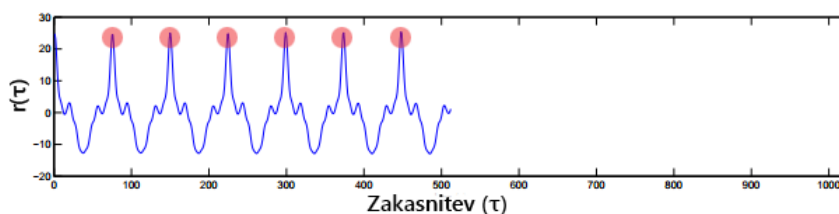
$$\text{osnovna frekvenca} = \frac{\text{frekvenca vzorčenja}}{\text{osnovna perioda}}. \quad (2.9)$$

Osnovna perioda je merjena v številu vzorcev, ki se pomnožijo s periodo vzorčenja. Perioda vzorčenja je inverzna vrednost frekvence vzorčenja.

Metoda temelji na opažanju, da ima zvok harmonične komponente, ki se v spektru ponavljajo z določeno periodo [7]. Pogosto je v rezultatih več maksimumov zaradi dobre korelacije na večkratnikih periode. Ti imajo lahko med seboj tudi zelo podobno višino. Zaradi tega je lahko izbrana napačna maksimalna vrednost in posledično napačna lokacija (slika 2.7).

Običajno se išče prvi maksimum. Iskanje je enostavnejše pri avtokorelaciji tipa 2, saj se tam vrednosti linearno zmanjšujejo proti ničli (slika 2.4). Tako imajo prednost začetne periode.

³Izračun avtokorelacije, pri katerem se vrednosti ne zmanjšujejo linearno.



Slika 2.7: Avtokorelacija z večkratniki periode s podobno višino [2].

Avtokorelacija v signalu uspešno odkriva povsem periodične segmente. Vendar pa v zvoku instrumenta ali petja segmenti včasih niso povsem periodični. Pogosto so prisotna nihanja, ki se kažejo kot variacije frekvence ali amplitude. Te variacije lahko vplivajo na uspešnost metode.

2.1.4 Središčno izrezovanje

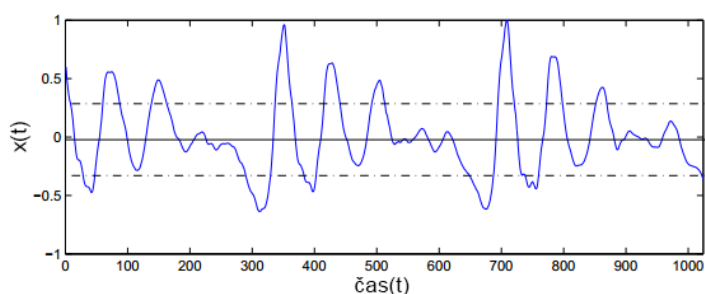
V signalu se večkrat nahajajo vrhovi, ki niso pomembni pri določanju maksimuma in lahko negativno vplivajo na njegov izbor. Vpliv teh vrhov lahko zmanjšamo s središčnim izrezovanjem⁴. Določiti moramo prag C_L , ki je definiran kot odstotek maksimalne amplitude signala.

$$y(t) = \begin{cases} 0 & \text{če je } |x(t)| \leq C_L \\ x(t) - C_L & \text{če je } x(t) > C_L \\ x(t) + C_L & \text{če je } x(t) < -C_L \end{cases} \quad (2.10)$$

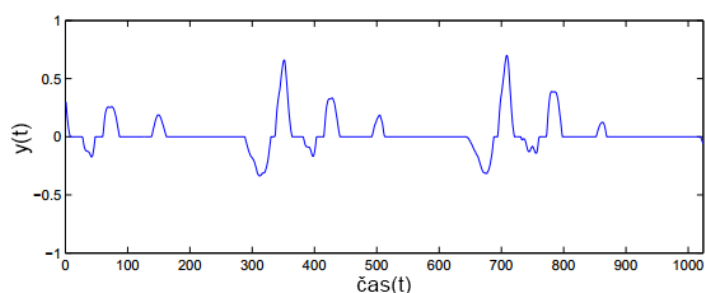
Glajenje signala z izrazom 2.10 poteka tako, da se poišče vzorce z absolutnimi vrednostmi pod pragom C_L . Tem vzorcem se pripiše vrednost 0. Ostalim vzorcem se prišteje ali odšteje vrednost pragu. Potem se njihove amplitude nahajajo na približno enaki ravni. Cilj glajenja je izničenje vpliva vokalnega trakta in ojačitev vokalnega vira.

Slika 2.8 prikazuje primer središčnega izrezovanja, kjer je za prag določena vrednost 30 %. Zglajen rezultat vsebuje bistveno manj nepomembnih vrhov.

⁴Angl. "Centre clipping". Prevod središčno izrezovanje je privzet iz [8].



(a) potek signala govora



(b) signal govora po središčnem izrezovanju

Slika 2.8: Primer središčnega izrezovanja na signalu govora [2].

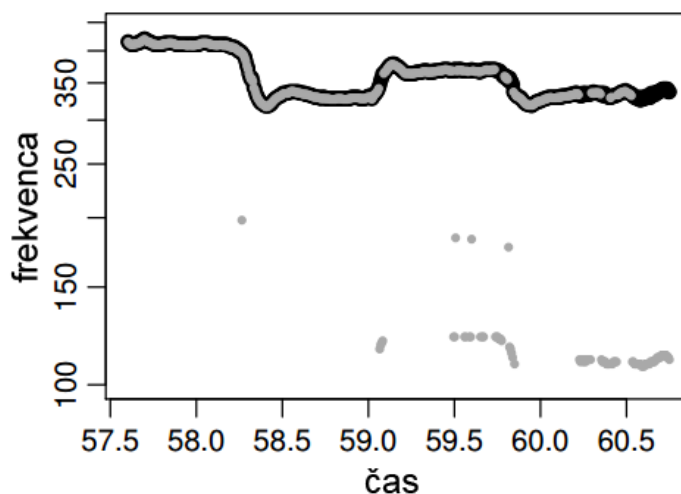
2.1.5 Metoda pYIN

V prvi polovici praktičnega dela smo uporabili Vamp vtičnik pYIN [6]. Metoda pYIN⁵ je izboljšava metode YIN. Obe temeljita na avtokorelaciji. Prva faza YIN metode je ocena osnovne frekvence za posamezen okvir. Ta ocena pogosto ni zanesljiva. Zato se v drugi fazi uporablja obdelava podatkov, s katero se dobljene rezultate filtrira. Filtriranje poveča uspešnost metode. Njena glavna pomanjkljivost je, da bi lahko bili v drugi fazi učinkovito uporabljeni še ostali kandidati, ki niso ohranjeni iz prve faze.

Metoda pYIN odpravlja to pomanjkljivost tako, da v prvi fazi izbere več kandidatov za osnovno frekvenco ter njihove verjetnosti. Ti so nato uporabljeni pri obdelavi in glajenju rezultatov v drugi fazi [9]. Slika 2.9

⁵Verjetnostni (angl. "Probabilistic") YIN.

prikazuje oceno frekvence obeh metod za štiri note. Izračuna sovpadata, kjer je siva barva obrobljena s črno barvo. Siva barva predstavlja frekvenčno sestavo izračunano z YIN metodo, črna barva pa s pYIN metodo. Med vrednostmi 59.5 in 61 na vodoravni osi so vidne številne napake YIN metode, ki jih metoda pYIN odpravlja.



Slika 2.9: Prikaz razlike med rezultati YIN in pYIN metode [9].

2.2 Intonacija

Od vsakega pevca se v pevski skupini brez instrumentalne spremljave zahteva zmožnost točnega petja. Poleg kvalitetne kontrole dihanja in izdelanega zvoka, je za natančno osnovno frekvenco pomembno tudi natančno poslušanje zvokov v okolici. Le tako se pevec lahko uskladi in poje točno s skupino [10].

Pri petju brez instrumentalne spremljave pevci velikokrat spremenijo intonacijo od začetka do konca skladbe. V literaturi so naštetih različni vzroki:

- težje ohranjanje intonacije pri nekaterih intervalih,
- razlika med naravno in temperirano uglasitvijo,

- avtomatsko nižanje intonacije pri skladbah napisanih v F-duru ali a-molu [11],
- izkušnost pevca,
- zahtevnost skladbe,
- psihološki stres [12] ipd.

Seaton v [11] raziskuje vzroke med pevci. Med pomembnimi vzroki izpostavi koncentracijo in utrujenost, s katerima je tesno povezan nadzor dihanja. Za izoblikovanje note je potrebno kontrolirano in enakomerno dovajanje zraka. Na to lahko zelo vpliva (fizična ali psihična) utrujenost pevca. Eden od dejavnikov je tudi samozavest. Če pevci niso prepričani v note, ki jih morajo odpeti, se to lahko pokaže v intonaciji. Pomanjkanje samozavesti je možno zaradi pomanjkanja vaje ali pevske nadarjenosti.

Premik v intonaciji si lahko razlagamo kot pojav, pri katerem pevec ali skupina pevcev začne peti skladbo v določenem tonu, zaključni ton pa je višji ali nižji od začetnega. Pojav je prisoten pri veliki večini pevcev [12].

Glasbeni instrumenti imajo tone določene s tipkami, strunami, prečkami ipd. Človeški glas nima tonov določenih na tak način, zato lahko pevec pri petju nehote določi preveliko ali premajhno višino med posameznimi toni.

Petje v skupini brez glasbene spremljave ponuja dodatne izzive za pevce. Posameznik lahko za svoj del v skladbi sam določi višino posamezne note, mora pa se prilagajati ostalim pevcem in skupaj z njimi tvoriti celoto.

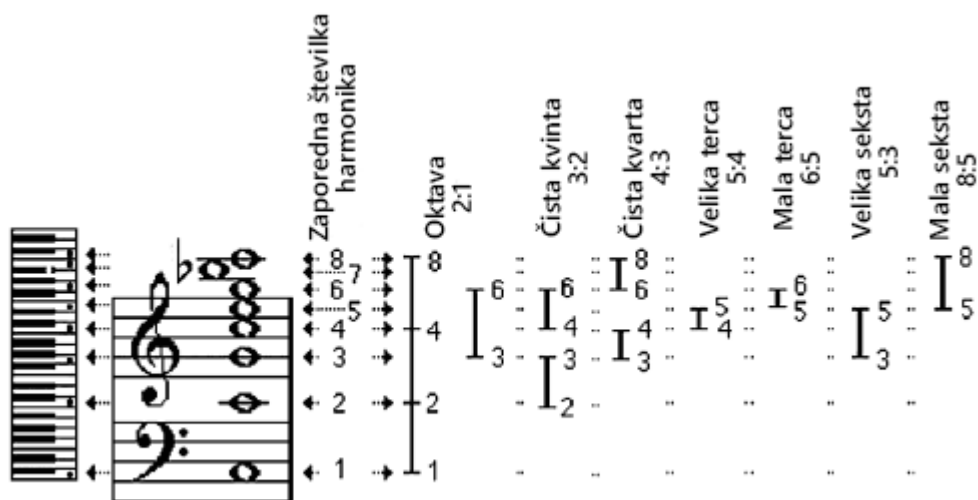
Problem določitve note je večji kot sprva zgleda, saj obstajata dva načina uglasitve, ki sta podrobneje opisana v nadaljevanju. Pri prvem zbor doseže čim bolj prijeten oziroma konsonanten⁶ rezultat. Pri drugem načinu pa ostaja v tonu⁷. Pri tem načinu zbor ne dosega konsonantnega rezultata [13].

⁶Konsonanca je obrazložena v naslednjem podpoglavju.

⁷Petje v tonu pomeni petje v točni intonaciji. Odpete note so točne in intonacija se ne spremeni od začetka do konca skladbe.

2.2.1 Naravna uglasitev

Pri doseganju konsonantnega zvoka naletimo na pojem naravna uglasitev⁸. Ko pevci pazljivo poslušajo dogajanje v glasbi, se bodo nagibali proti naravni uglasitvi. V taki uglasitvi so posamezni intervali⁹ določeni z razmerji naravnih števil. Ta razmerja nastanejo med komponentami naravne harmonične vrste. Prvih osem harmonikov tona C3 je prikazanih na sliki 2.10. Na sliki so zapisani tudi intervali med njimi in njihova frekvenčna razmerja. Tako uglašen akord¹⁰ predstavlja popolno harmonijo.



Slika 2.10: Prvih osem harmonikov v naravni harmonični vrsti [13].

Ton, ki ga zapoje pevec vsebuje veliko harmonikov. Njihova amplituda je odvisna od formantnih frekvenc¹¹ in kvalitete glasu. Ko skupaj pojeta dva pevca, bo vsak harmonik prvega pevca imel najbližji sosednji harmonik iz množice harmonikov drugega pevca. Takrat ko sta osnovni frekvenci obeh pevcev uglašeni v razmerju naravnih števil (npr. 3:2), bodo sovpa-

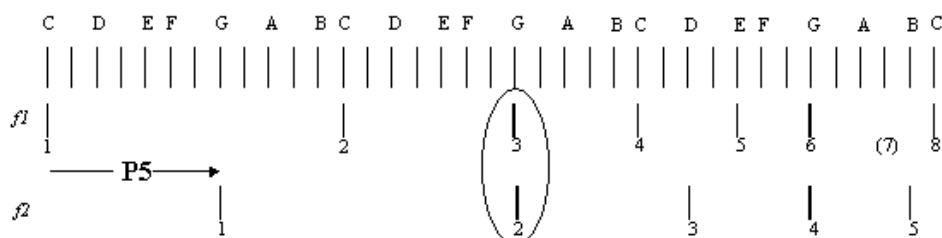
⁸Angl. "Just intonation".

⁹Interval: razmik med tonoma.

¹⁰Akord je sozvočje najmanj treh tonov.

¹¹Resonančne frekvence govornega trakta pri izreku določenega glasu ali petju.

dali tretji harmonik prvega pevca in drugi harmonik drugega pevca ter njuni večkratniki. Tak primer prikazuje slika 2.11.



Slika 2.11: Sovpadanje harmonikov pri razmerju 3:2 [14].

Ko harmoniki sovpadajo – so čisto uglašeni, se akord dojema kot konsonanten, v nasprotnem primeru pa kot disonanten. Različna stopnja konsonance ali disonance med sosednjimi harmoniki prispeva k splošnemu dojemanju akorda. Stopnja konsonance prispeva k občutku prijetno poslušljivega akorda. Nasprotno pa stopnja disonance prispeva k občutku neprijetno poslušljivega akorda. Največjo stopnjo konsonance bodo imeli akordi uglašeni v naravni uglasitvi. V tej uglasitvi so toni v akordu določeni z razmerji naravnih števil¹². Število popolnoma sovpadajočih sosednjih harmonikov je takrat največje [13].

2.2.2 Temperirana uglasitev

Žal ni mogoče ohranjati naravne uglasitve med vsemi toni na enem instrumentu. Zakaj je temu tako, lahko prikažemo matematično. Recimo, da bi želeli sami uglasiti klavir. Izberemo si najnižjo A tipko in ji določimo frekvenco 27.5 Hz ter ime A_0 ¹³. Od te note se začnemo pomikati za čisto kvinto navzgor. V tabeli 2.1 najdemo podatek, da čista kvinta znaša 7 poltonov (7 tipk na klavirju). Novo tipko uglasimo z razmerjem 3:2. Podatek o razmerju

¹²Npr. poljubno izbrani prvi ton, drugi ton uglašen v razmerju 3:2, tretji pa v razmerju 5:4.

¹³Po standardu ima najnižja A tipka na klavirju frekvenco 27.5 Hz.

se nahaja na sliki 2.10 - razmerje tretjega harmonika. Vrednost 27.5 Hz pomnožimo s tem razmerjem in dobimo 41.25 Hz.

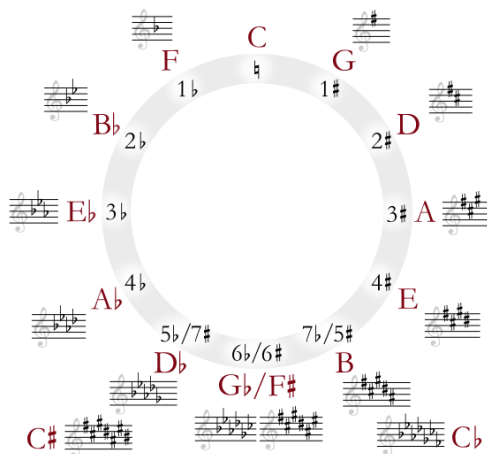
Interval	Razlika v poltonih
Prima	0
Mala sekunda	1
Velika sekunda	2
Mala terca	3
Velika terca	4
Čista kvarta	5
Zmanjšana kvinta	6
Čista kvinta	7
Mala seksta	8
Velika seksta	9
Mala septima	10
Velika septima	11
Oktava	12

Tabela 2.1: Razlika v poltonih med različnimi intervali.

Take premike ponavljamo, dokler se po dvanajstih korakih ponovno ne znajdemo na tipki A . Poimenujemo jo A_7 , saj je od prvotne note A_0 oddaljena 7 oktav¹⁴. Izračunamo ji frekvenco 3568.02 Hz. Posamezni koraki pri pomikanju so vidni na sliki 2.12, pojav se imenuje kvintni krog. S pomočjo tega pojava bi lahko sklepali, da je mogoče v naravni intonaciji uglasiti vse tipke na klavirju. V kvintnem krogu se nahajajo vsi toni, ostale tipke bi lahko izračunali s pomočjo oktave in razmerja 2:1 (enaka tipka, le za 12 poltonov višje ali nižje).

Pred tem preverimo izračun frekvence note A_7 . Frekvenci note A_0 sedemkrat pomnožimo razmerje oktave 2:1. V tem primeru znaša frekvenca note A_7 3520 Hz.

¹⁴Dvanajst premikov za sedem poltonov ali sedem premikov za dvanajst poltonov (oktava).



Slika 2.12: Kvintni krog [15].

Ob primerjavi obeh rezultatov opazimo, da se razlikujeta. Razmerje med njima je znano pod imenom pitagorejska razlika¹⁵ in znaša 1.01364 [13]. Prikaže ga naslednji izraz:

$$\frac{\left[\frac{3}{2}\right]^{12} \times 27.5 \text{ Hz}}{[2]^7 \times 27.5 \text{ Hz}} = \frac{129.746 \times 27.5 \text{ Hz}}{128.000 \times 27.5 \text{ Hz}} = \frac{3568.02 \text{ Hz}}{3520 \text{ Hz}} = 1.01364. \quad (2.11)$$

Pri sedmih oktavah razlika znaša 48.02 Hz. Zaradi te razlike na klavirju ne moremo hkrati ohranjati naravno uglasenih čistih kvint (3:2) in oktav (2:1).

Do 19. stoletja se je uporabljal način uglasovanja, pri katerem je bilo na instrument mogoče izvajati le skladbe v določenih lestvicah. V tem načinu so bile oktave uglasene v naravni uglasitvi. Za tiste lestvice v katerih je bila na instrument izvajana glasba, so bili ostali intervali čim bolj približani naravni uglasitvi. Za ostale tone so bile vrednosti izračunane z razmerjem pitagorejske razlike. Te tone so uporabljani zelo redko ali pa skoraj nikoli, ker so zveneli zelo disonantno.

¹⁵Angl. "Pythagorean comma".

Interval	Naravna uglasitev	Temperirana uglasitev
Prima*	$1:1 = 1.000$	1.000
Mala sekunda°	$25:24 = 1.042$	1.059
Velika sekunda°	$9:8 = 1.125$	1.122
Mala terca*	$6:5 = 1.200$	1.189
Velika terca*	$5:4 = 1.250$	1.260
Čista kvarta*	$4:3 = 1.333$	1.335
Zmanjšana kvinta°	$44:32 = 1.406$	1.414
Čista kvinta*	$3:2 = 1.500$	1.498
Mala seksta*	$8:5 = 1.600$	1.587
Velika seksta*	$5:3 = 1.667$	1.682
Mala septima°	$9:5 = 1.800$	1.782
Velika septima°	$15:8 = 1.875$	1.888
Oktava*	$2:1 = 2.000$	2.000

Tabela 2.2: Razlika v razmerjih med naravno in temperirano uglasitvijo (* konsonant, °disonant) [13, 16, 17].

Recimo, da imamo klavir, ki je uglašen tako, da so v naravni uglasitvi uglašene tipke, ki se pojavijo v lestvici C-dura¹⁶. Skladba zaigrana v C-duru bo zvenela konsonantno, saj so vsi intervali uglašeni v naravni intonaciji. Če na tako uglašen klavir zaigramo skladbo v D-duru, bo zvenela disonantno. Tipke, ki se nahajajo v D-duru ne bodo uglašene v naravni intonaciji.

V 19. stoletju so izdelovalci začeli izdelovati klavirje in orgle na katerih so bili vsi poltoni med oktavami uglašeni s kompromisom. Tako uglasitev poznamo pod imenom temperirana uglasitev¹⁷. Pri njej je razlika med vsemi poltoni enaka. Razmerje oktave (2:1) se enakomerno razdeli med 12 poltonov z naslednjim izrazom:

¹⁶Durova lestvica: zaporedje osmih tonov. Osmi ton je enak prvemu. Večina tonov se od prejšnjega razlikuje za dva poltona, izjema sta četrti in osmi ton, ki se od prejšnjega razlikujeta za en polton.

¹⁷Angl. "Equal temperament".

$$\sqrt[12]{2} = 1.0595. \quad (2.12)$$

Tako je bilo zadoščeno želji skladateljev in izvajalcev - igranje skladb v poljubni lestvici. Vendar v taki uglasitvi noben od intervalov (razen oktave) ni točno uglasen in vsebuje določeno stopnjo disonance. Dandanes so se ljudje že privadili na tako uglasitev in ko slišijo glasbo v naravni uglasitvi je učinek opisan kot izjemen [13].

Nota	Naravna uglasitev	Temperirana uglasitev	Razlika
A_4	440.000	440.000	0.000
$A\#_4$	458.334	466.164	-7.830
H_4	495.000	493.883	1.117
C_5	528.000	523.251	4.749
$C\#_5$	550.000	554.365	-4.365
D_5	586.667	587.330	-0.663
$D\#_5$	605.000	622.254	-17.254
E_5	660.000	659.255	0.745
F_5	704.000	698.456	5.544
$F\#_5$	733.333	739.989	-6.655
G_5	792.000	783.991	8.009
$G\#_5$	825.000	830.609	-5.609
A_5	880.000	880.000	0.000

Tabela 2.3: Razlika v Hz med frekvencami naravne in temperirane uglasitve [18].

Pevec lahko izbere v kateri uglasitvi bo pel, saj lahko prosto določi višino tona. Zaradi poslušanja glasbe in vaje s klavirjem naj bi se posamezen pevec nagibal k temperirani uglasitvi. Ko pa v skupini več pevcev poje note, ki v večglasju sestavljajo akorde, se pri pazljivem poslušanju pojavi tendenca proti naravni uglasitvi [10]. Razlika med intervali pri naravni in temperirani uglasitvi so prikazane v tabeli 2.2. Razlike v Hz med posameznimi toni pa v tabeli 2.3.

2.2.3 Problem točne intonacije

Pevci, trobila in instrumenti brez prečk pogosto proizvajajo zvoke, ki zaradi različnih uglasitev niso v tonu z ostalimi instrumenti (npr. s klavirjem). Te razlike so zelo majhne, vendar so bistvenega pomena pri ustvarjanju glasbe na visokem nivoju. Iz tega sledi nov problem. Ko je skladba sestavljena iz različnih lestvic, lahko pri petju te razlike v uglasitvi posameznih not povzročijo premike v intonaciji. Kot posledica sta prvi in zadnji akord različno uglašena. Spreminjanje intonacije je opaženo. Nekateri zmotno mislijo, da so za to krivi pevci. Dejansko pa lahko pevci pojejo kvalitetno in ustvarjajo zelo konsonanten zvok, pri čemer je premik v intonaciji neizogiben [10].

Howard v [13] potrdi hipotezo: pri štiriglasnem petju vokalnega kvarteta brez glasbene spremljave bo pri modulaciji akordov prišlo do premikov v intonaciji, če se bodo pevci nagibali k ne temperirani uglasitvi.

Z različnimi pevskimi vajami je bilo dokazano, da je pri petju v naravni uglasitvi intonacija pevcev drsela navzdol. Vredno je omeniti, da je do precejšnjih premikov v intonaciji prišlo tudi pri vajah, ki so bile kratke in so vsebovale majhno število akordov. Pokazalo se je, da pevci v zborovskem sestavu brez instrumentalne spremljave raje pojejo v naravni uglasitvi.

Howard v [13] še predpostavi, da je premik v intonaciji nujen, če je skladba sestavljena iz več lestvic in želijo pevci peti v naravni uglasitvi. Pevec je postavljen pred odločitev, naj poje v tonu ali naj dovoli premike v intonaciji. Pevska skupina brez spremljave torej mora zaiti v intonaciji, če hoče ostati v tonu od začetka do konca skladbe.

Pevci lahko spremenijo njihovo intonacijo podzavestno. K petju v naravni uglasitvi je mogoče dodatno prispevati tako, da se na začetku skladbe namesto prvega akorda zaigra samo eno noto na klavirju. To bo izboljšalo tudi poslušanje. Odsvetovana je vaja ob klavirju, ki je uglašen v temperirani uglasitvi. Odsvetovano je tudi zaigrati celoten akord pred začetkom izvajanja skladbe.

Kako sploh pevci lahko določijo višino intervala, ki bi bil skladen z naravno uglasitvijo? Howard v [10] zapiše, da verjetno obstaja psiho-akustična

osnova za stopnjo disonance v zvoku.

Za premike v intonaciji ne gre kriviti pevcev, saj lahko le s premiki ustvarjajo konsonanten zvok. Če je potrebna korekcija intonacije, je to potrebno doseči nekako drugače, npr. z določitvijo točk v skladbi kjer pride do premikov v intonaciji. Kot drugo možnost se omenja še način petja, kjer bi pevec ali skupina pevcev z glavno melodijo peli v temperirani uglasitvi. Ostali, ki s svojimi toni postavljajo konsonantne harmonije v naravni uglasitvi, pa bi se med skladbo prilagajali glavni melodiji. Kar bi zagotovilo prijeten zvok v akordih in točnost v intonaciji. To pa bi za pevca glavne melodije predstavljajo zelo velik izziv [10].

Poglavje 3

Uporabljena programska orodja

V tem poglavju so opisana programska orodja, ki smo jih uporabili v praktičnem delu diplomske naloge. Za delno obdelavo in analizo posnetkov smo uporabili Audacity [19] in Sonic Visualiser [20]. Za preostanek analize posnetkov in obdelavo ter prikaz rezultatov pa smo uporabili programski jezik Python in njegove pakete.

3.1 Programski jezik Python

Python je ustvaril Guido Van Rossum v začetku devetdesetih let prejšnjega stoletja. Ime je dobil po nanizanki Leteči cirkus Montya Pythona. Spada med skriptne programske jezike in se ga lahko uporablja za najrazličnejše programske naloge. Odličen je za učenje programiranja pri začetnikih. V Google in Yahoo! ga uporabljajo pri iskalnikih in diskusijskih skupinah, v NASI pa za programske sisteme. V veliki meri se uporablja za sistemsko administracijo, v njem so napisane komponente LINUX distribucij.

Python se avtomatsko prevaja v binarno kodo, nato se običajno shrani na disk. Ob ponovnem zaganjanju ni več potrebno prevajanje, razen če pride do sprememb v izvorni kodi. Python ima dinamične podatkovne tipe [21]. Značilnost Pythona je pomembnost zamikov v kodi, namesto uporabe oklepajev. Prav tako pa tudi dostopnost za vse platforme. Lahko deluje na opera-

cijskem sistemu Microsoft Windows, Macintosh in vseh distribucijah Linuxa. Zaradi tega je zelo prenosen. Če je napisan za eno platformo, se enostavno uporablja tudi na drugih. Zagotavlja zmogljiv asortiment vgrajenih tipov (seznami, slovarji, nizi) ter veliko vgrajenih funkcij in nekaj konstruktorjev. Vsebuje tudi standardno knjižnico, ki vključuje več modulov za široko paleto storitev, vključno z regularnimi izrazi in TCP/IP sejami.

3.2 Paketi in knjižnice

V tem podpoglavju se nahajajo paketi in knjižnice, ki smo jih uporabili pri implementaciji sistema za analizo posnetkov. Opisane so tudi njihove glavne značilnosti.

3.2.1 Audiolab

Audiolab je Pythonov paket za delo z avdio datotekami. Podpira različne avdio formate, kot so: wav, aiff, au, flac, ogg in htk [22].

3.2.2 NumPy

NumPy je temeljni paket za numerično obdelavo podatkov v Pythonu. Med drugim vsebuje [23]:

- zmogljive N-dimenzionalne objektne sezname,
- visoko zmogljive matematične funkcije,
- orodja za integracijo C/C++ in Fortran kode,
- linearno algebro, Fourierjevo transformacijo in naključna števila.

Uporaba paketa NumPy skupaj s Pythonovim naborom funkcij omogoča funkcionalnosti, ki so kompatibilne z Matlabom. Na voljo so tudi dopolnilni paketi. SciPy je knjižnica, ki dodaja več funkcionalnosti, podobnih kot v Matlabu. Matplotlib pa je paket za risanje, ki prav tako omogoča funkcionalnosti, ki so kompatibilne z Matlabom.

3.2.3 SciPy

SciPy je odprtokodna Pythonova knjižnica za znanstveno in tehnično obdelavo podatkov [24]. Vsebuje module za:

- optimizacijo,
- linearno algebro,
- integracijo,
- interpolacijo,
- posebne funkcije,
- Fourierjevo transformacijo,
- obdelavo slik in signalov,
- reševanje diferencialnih enačb in ostalih nalog.

Temelji na NumPyevem objektu in je del zbirke NumPy [25].

3.2.4 Matplotlib

Pythonova knjižnica Matplotlib avtorja Johna Hunterja (1968-2012) omogoča izris dvodimenzionalnih grafov. Kakovost izrisa ustreza publikacijam, bodisi v tiskani obliki ali v interaktivnih okoljih. Lahko se jo uporablja v Pythonovih skriptah, lupinah, spletnih aplikacijskih strežnikih in orodjih grafičnega uporabniškega vmesnika.

Odprtokodna knjižnica poskuša enostavne stvari ohraniti enostavne, težke pa poenostaviti. Omogoča generiranje raznoraznih vrst grafov z le nekaj vrsticami kode. Enostaven izris omogoča vmesnik, ki spominja na Matlabove vmesnike. Uporabnik ima preko objektno usmerjenega vmesnika nadzor nad stilom linije, pisavami, lastnostmi osi ipd. Obstaja več dodatnih orodij, med drugim tudi tridimenzionalni izris [26].

3.3 Audacity

Audacity je odprtokodna programska oprema za snemanje in urejanje zvoka. Na voljo je za Windows, OS X, Linux in druge operacijske sisteme. Aplikacijo sta začela razvijati Dominic Mazzoni in Roger Dannenberg leta 1999, izdana je bila v letu 2000 [27].

Poleg snemanja iz različnih virov omogoča tudi naknadno obdelavo vseh vrst zvoka. Obdelava vključuje dodajanje učinkov, kot so: normalizacija, obrezovanje, presihanje ipd [19]. Za potrebe te naloge so bile uporabljene naslednje funkcije: predvajanje zvoka, uvoz in izvoz wav datotek, rezanje posnetka, normalizacija, ojačitev signala in odstranitev nizkih amplitud.

3.4 Sonic Visualiser

Sonic Visualiser je brezplačna programska oprema, ki omogoča vizualizacijo, analizo in anotacijo zvočnih datotek. Njen razvoj je potekal na londonski Univerzi v centru za digitalno glasbo. Program se uporablja v glasbi pa tudi v znanstvene namene. Deluje na platformi Linux, OS X in operacijskem sistemu Windows [20]. Omogoča uporabo visoko specializiranih vtičnikov formata Vamp. V nalogi je bil uporabljen vtičnik pYIN [6]. Rezultati analiz so bili izvoženi v datoteke csv¹ oblike.

¹Angl. "Comma-separated values".

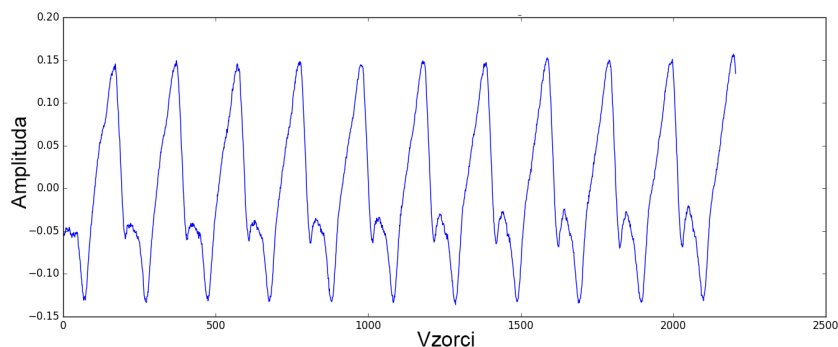
Poglavje 4

Implementacija sistema za analizo posnetkov

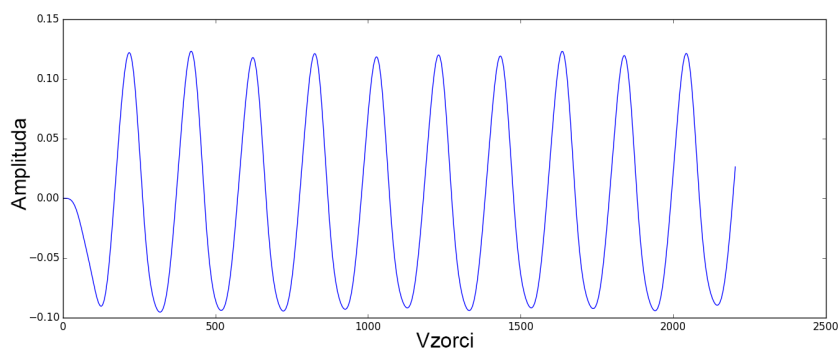
V tem poglavju je predstavljena implementacija programov za analizo petja. Za potrebe diplomske naloge smo napisali tri programe: *AnalizaPosnetkov.py*, *PrikazRezultatov1.py* in *PrikazRezultatov2.py*¹. Izbrali smo programski jezik Python zaradi njegove enostavnosti in učinkovitosti. Prvi program deluje tako, da analizira zvočne datoteke wav formata. Posamezen posnetek je s paketom Audiolab pretvorjen v seznam vrednosti med -1 in 1. Seznam se razdeli v okvirje, ki se posamezno obdelajo. Dolžina okvirja je določena v mili sekundah. Najboljši rezultati so bili opaženi pri dolžini 35 milisekund.

Najprej se s pomočjo knjižnice SciPy nad okvirjem izvede nizko propustni filter. V tem koraku se odstranijo vse frekvenčne komponente, ki bi lahko negativno vplivale na izračun osnovne frekvence in niso pomembne pri sledečih izračunih. Ker smo vedeli, kakšni posnetki bodo analizirani in smo priskrbeli notni zapis, smo lahko z gotovostjo določiti zgornjo mejo, nad katero se ne pojavi nobena osnovna frekvenca. Ta meja je bila določena dva poltona nad najvišjo noto, ki jo je zapel najvišji ženski glas – sopran. Razlika pred in po nizkoprepustnem filtriranju je vidna na slikah 4.1 in 4.2.

¹<https://github.com/martins90/AnalizaPetja.git>



Slika 4.1: Prikaz vzorcev samoglasnika "a".



Slika 4.2: Vzorci samoglasnika "a" po nizkoprepustnem filtriranju.

4.1 Določanje osnovne frekvence

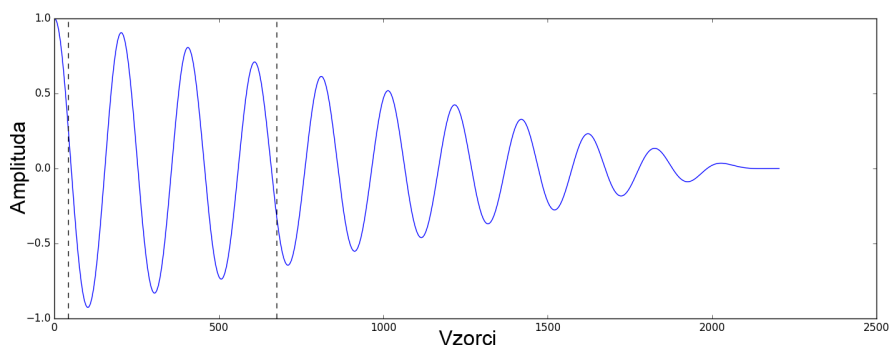
Za določanje osnovne frekvence je bila izbrana metoda avtokorelacije. Pri implementaciji smo uporabili prosto dostopen algoritem pod MIT licenco [28], ki smo ga prepisali iz programskega jezika C++ v Python. Narejene so bile tudi nekatere dodatne spremembe. Kot avtor v [28] svetuje, je algoritem hitrejši z uporabo Fourierjeve transformacije namesto križne korelacije. Avtokorelacija signala je enaka inverzni Fourierjevi transformaciji produkta Fourierjeve transformacije signala in njene konjugirane vrednosti. To enakost prikazuje naslednji izraz:

$$AC(x) = \text{IFFT}(\text{FFT}(x)\text{FFT}(x)^*). \quad (4.1)$$

Sledi iskanje najvišje vrednosti. V tem koraku avtor algoritma v [28] izvaja iskanje vrednosti med najmanjšo in največjo pričakovano periodo. Ti dve periodi sta izračunani iz najvišje in najnižje pričakovane frekvence z naslednjim izrazom:

$$\text{minimalna perioda} = \frac{\text{frekvenca vzorčenja}}{\text{maksimalna frekvenca} - 1}. \quad (4.2)$$

Slika 4.3 prikazuje rezultat avtokorelacije tipa 2 iz podpoglavja 2.1.3. Prekinjene navpične črte označujejo mejnike od kje do kje poteka iskanje najvišje vrednosti.



Slika 4.3: Avtokorelacija samoglasnika "a".

Po določeni najvišji vrednosti se preveri, če mogoče ta najvišja vrednost ni vrh korelacije. Tak primer se zgodi, ko se vrh nahaja izven določenih mejnih vrednosti period. Pred končnim določanjem osnovne periode v okvirju, se vrh poišče z interpolacijo. Če je desna vrednost višja od leve vrednosti, se vrh nahaja med trenutno in desno pozicijo.

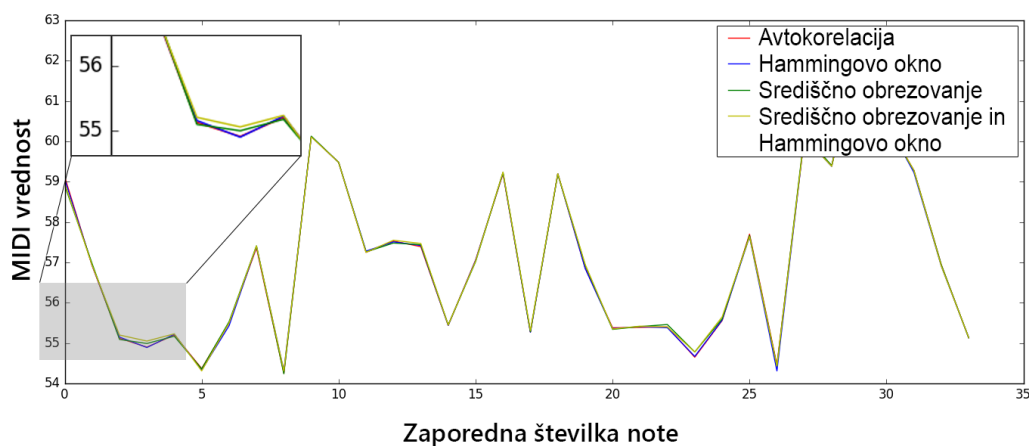
Ko je osnovna perioda (izražena v številu vzorcev) določena, je iz nje mogoče izračunati osnovno frekvenco z naslednjim izrazom:

$$\text{osnovna frekvenca} = \frac{\text{frekvenca vzorčenja}}{\text{osnovna perioda}}. \quad (4.3)$$

Vrednosti smo pri analizi zaokrožili na dve decimalni mesti.

V poglavju 2 so omenjeni različni načini, ki bi lahko pripomogli k izboljšavi avtokorelacije (Hammingovo okno, središčno izrezovanje). Na testnem posnetku poljubne ljudske pesmi so bili preizkušeni štirje različni načini:

- osnovni izračun avtokorelacije,
- izračun avtokorelacije okvirja, pomnoženega s Hammingovim oknom,
- izračun avtokorelacije okvirja, zglajenega s središčnim izrezovanjem,
- izračun avtokorelacije okvirja, zglajenega s središčnim izrezovanjem in pomnoženega s Hammingovim oknom.



Slika 4.4: Razlika pri različnih izračunih avtokorelacije.

Na sliki 4.4 je vidno, da sicer razlika pri izračunanih vrednostih obstaja, vendar se ta pojavi le pri nekaterih notah (3, 12, 23). Za note, kjer je bila

razlika očitna, smo osnovno frekvenco preverili s programom Sonic Visualiser. Tudi tam so se rezultati razlikovali med različnimi metodami za izračun osnovne frekvence. Sklenili smo, da bomo pri obdelavi posnetkov uporabili osnovni izračun avtokorelacije, kar ne bo povečevalo časovne zahtevnosti našega programa. Majhne razlike v izračunu osnovnih frekvenc za posamezno noto pa bodo zanemarljive.

Izračunane frekvence so nato pretvorjene v prijaznejši zapis. Na tem mestu je bil uporabljen MIDI protokol, ki je natančneje opisan v naslednjem podpoglavju.

Za filtriranje dobljenih rezultatov uporabimo Algoritem 1, ki se nahaja na naslednji strani. Napisali smo ga samo za ta konkretni problem. Algoritem izloči morebitne vrednosti not, ki izstopajo. Za njih se lahko z gotovostjo sklepa, da so bodisi šum v posnetku, bodisi konsonant pri petju. Njegove frekvenčne komponente so drugačne kot pri samoglasnikih. Take izstopajoče vrednosti so ponavadi na začetku ali koncu note (s, š, t, k ipd). Ugotavljamo jih na podlagi stopnje odstopanja od ostalih vrednosti pri trenutni noti.

Podatki so po filtriranju zglajeni in pripravljeni za nadaljnjo obdelavo. Sledi združevanje not in računanje mediane z metodo prikazano na sliki 4.5. Domislili smo se načina, kako note združiti avtomatsko. Za pravilno delovanje metode je potrebno melodijo odpeti s presledki med notami. Metoda se z zanko sprehaja čez vrednosti in jih dodaja v ločen seznam imenovan "kandidati". Ko se v podatkih pojavi ničelna vrednost (presledek med notami), se za kandidate izračuna mediana. Ta mediana je dodana v seznam not in predstavlja vrednost trenutne note. Nato se izprazni seznam kandidatov in nadaljuje s pregledom v zanki. Metoda hkrati preverja, če so se med kandidati znašle take vrednosti, ki ne predstavljajo note (premalo kandidatov). Take vrednosti metoda ne obravnava kot ton in jih zanemari.

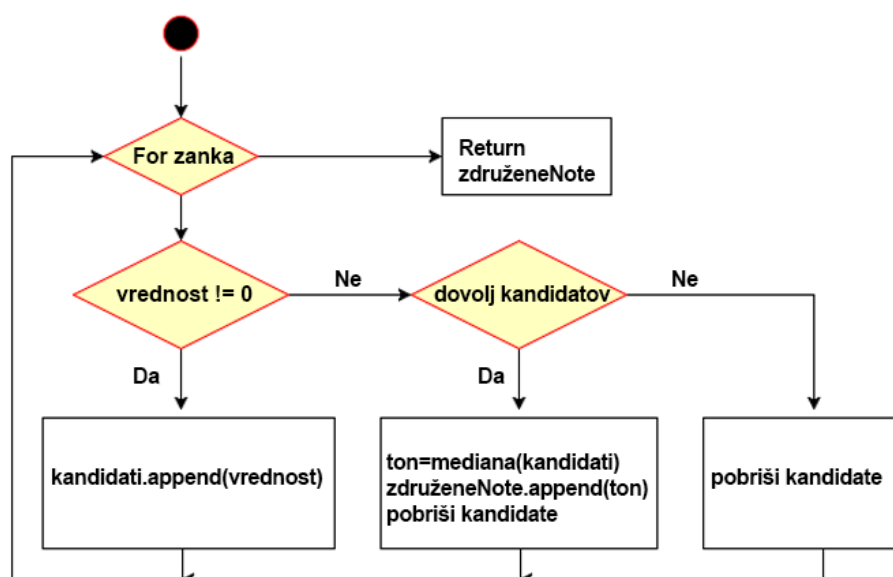
Ob koncu postopka dobimo seznam not, ki so bile posnete pri petju.

Algoritem 1 Filtriranje rezultatov

```

1: procedure FILTRIRANJE(seznam, razlika)
2:    $n \leftarrow 0, vrednosti \leftarrow 0$ 
3:   for vrednost in seznam do
4:     if vrednost  $\neq 0$  then
5:       if  $n == 0$  then ▷ Prva vrednost, ki ni 0 (nova nota)
6:          $vrednosti \leftarrow vrednost, n \leftarrow 1$ 
7:       else
8:         if  $abs(vrednost - vrednosti/n) > razlika$  then ▷ Razlika
           med vrednostjo in povprečjem ostalih je prevelika
9:           if prejsnjaVrednostIzstopa() then
10:             $prejsnjavrednost \leftarrow 0, n \leftarrow 1$ 
11:             $vrednosti \leftarrow vrednost$ 
12:          else
13:             $n \leftarrow 0, vrednost \leftarrow 0, vrednosti \leftarrow 0$ 
14:          end if
15:        else
16:           $n \leftarrow n + 1$ 
17:           $vrednosti \leftarrow vrednosti + vrednost$ 
18:        end if
19:      end if
20:    else
21:      if  $n < 2$  then ▷ Premalo uporabnega signala
22:         $seznam \leftarrow$  pobriši zadnjih  $n$  vrednosti
23:      end if
24:       $n \leftarrow 0, vrednosti \leftarrow 0$ 
25:    end if
26:  end for
27:  return seznam
28: end procedure

```



Slika 4.5: Delovanje metode za združevanje not.

4.2 Notni zapis MIDI

MIDI je protokol, ki določa oštevilčenje not. Vsak polton je predstavljen z eno enoto (številko). Številko note se izračuna z naslednjim izrazom:

$$n = 69 + 12 \log_2 \left(\frac{f}{440} \right), \quad (4.4)$$

kjer je f frekvenca, n številka note, število 69 pa predstavlja srednji A ton, katerega frekvenca je 440 Hz. Številka note je lahko tudi realno število v primeru, ko določen ton odstopa od referenčnega tona za manj kot pol tona [2]. Primer: srednji C ton ima frekvenco 261.63 Hz in je v MIDI lestvici oštevilčen kot 60. Ton izmerjen pri 260 Hz bi bil oštevilčen kot 59.89. Razlika med toni torej znaša 0.11 poltona oziroma 11 stotin [1]. Ker so razlike v tonih na splošno majhne, so note zaokrožene na dve decimalni mesti.

4.3 Prikaz rezultatov

Za prikaz rezultatov smo implementirali 2 programa. Prvi prikaže rezultate analize pevcev, ki so samostojno odpeli melodijo. Drugi prikaže rezultate štirih pevcev, ki so sočasno odpeli pevsko vajo.

PrikazRezultatov1.py prebere csv datoteke, ki smo jih izvozili v Sonic Visualiserju. Sledi primerjava prebranih not z notnim zapisom. Za notni zapis smo ročno napisali zaporedje celih števil², ki predstavlja razliko med sosednjimi notami. Razlika je izražena v poltonih. S pomočjo tega zaporedja so izračunane napake pri petju melodije. Napake predstavljajo razlike med dejanskimi in pričakovanimi notami. Vse napake se shranijo v slovar. V njem ključ predstavlja zaporedno številko note. Zadnji korak je izris napak po notah ter izris skupne mediane od začetka do konca skladbe (slika 5.3 v naslednjem poglavju).

Program *PrikazRezultatov2.py* prebere podatke, ki jih pridobimo po izvajanju programa *AnalizaPosnetkov.py*. Dva različna programa za analizo smo morali napisati, ker se pri posnetkih kvarteta notni zapis med glasovi razlikuje. Prva skripta ni bila uporabna in jo je bilo potrebno nekoliko spremeniti. Namesto zaporedja razlik med sosednjimi toni, smo v tem primeru notni zapis ročno pretvorili v MIDI oštevilčenje. Za prvo noto smo izračunali MIDI enoto. Za vsako naslednjo noto smo od prejšnje odšteli ali prišteli število poltonov³.

Podobno kot v prejšnjem programu, ki prikazuje rezultate, se od dejanskih vrednosti odšteje MIDI zapis skladbe. Razlike, ki prikazujejo napake, se izrišejo v grafu. Te razlike prikazuje npr. slika 5.7 v podpoglavju 5.2.

V naslednjem poglavju se nahaja razlaga rezultatov.

²Zaporedje skladbe Veseli pastir: 0,9,-4,-3,-2,2,-5,3,-2,2,10,-3,-2,-1,-2,-2,9,-9,9,-4,-3, -2,2,-5,3,-2,2,10,-3,-2,-1,-2,2,1.

³Ton A je oštevilčen kot 69. Ton C je tri poltone nad tonom A, torej je oštevilčen z $69+3=72$.

Poglavje 5

Rezultati analize realnih posnetkov

Pri analizi in snemanju je sodelovalo 23 pevcev Akademskega pevskega zbora Univerze na Primorskem. Med pevci je bilo približno enako razmerje v višini glasov (sopran, alt, tenor, bas). Analiza je bila izvedena v dveh delih, ki sta predstavljena v nadaljevanju.

5.1 Iskanje napak pri intervalih

V prvem delu smo se posvetili iskanju intervalov. Premik v intonaciji bi se lahko zgodil na mestih, kjer pevci ne bi bili točni v določenih intervalih. Za analizo sta bili izbrani dve skladbi: ljudska Tam dol na ravnem polju (slika 5.1) in Veseli pastir (slika 5.2) Simona Gregorčiča, poznana tudi pod imenom Zakrivljeno palico v roki. Prva je bila izbrana zaradi ponavljajočega se intervala male terce¹. Druga pa zaradi petih velikih skokov intervala velike sekste in male septime².

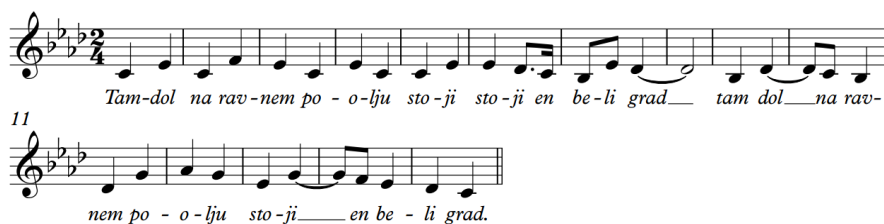
¹V skladbi Tam dol na ravnem polju se mala terca (navzgor ali navzdol) pojavlja v 1., 3., 4., 5. in 9. taktu, prav tako pa tudi med 10. in 11. taktom.

²Veliki skoki se pojavijo pri prehodu iz 1. v 2., 5. v 6., 7. v 8., 9. v 10. in 13. v 14. takt.

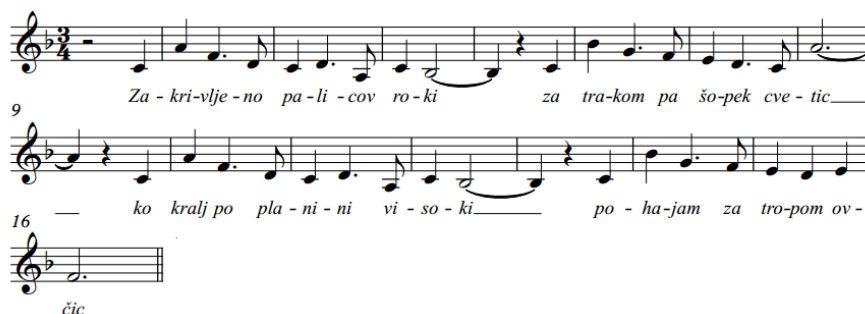
Notni material smo napisali za visoke in nizke glasove v F in B oziroma As in D-duru. Začetni ton je bil zaigran na klavir.

Značilnosti postopka zajema petja:

- uporaba Shure SM 58 vokalnega mikrofona,
- zvok je bil posnet s snemalno napravo Zoom H5,
- frekvenca vzorčenja 44100 Hz,
- 16-bitna resolucija,
- obdelava v programu Audacity (normalizacija ter rezanje začetka in konca).



Slika 5.1: Notni zapis skladbe Tam dol na ravnem polju.



Slika 5.2: Notni zapis skladbe Veseli pastir.

Izkazalo se je, da je avtomatska analiza petja z besedilom večji izziv, kot je bilo sprva pričakovano. Ob testiranju programa *AnalizaPosnetkov.py* se je ta izkazal kot učinkovit. Ob množici različnih pevcev pa ni zagotavljal pravih rezultatov. Za pravilno delovanje avtomatske analize program zahteva izrazito ločene tone s presledki vsaj 35 milisekund. Kljub opozorilu pri snemanju, so pevci pri nekaterih notah naredili premajhne razmike, ali pa so jih celo izpustili. Pri takih posnetkih delovanje programa ni bilo pravilno.

Sprejeta je bila odločitev, da bodo z implementirano rešitvijo analizirani posnetki kvarteta v drugem delu. Posnetki pevcev pa bodo obdelani v orodju Sonic Visualiser z vtičnikom pYIN algoritma.

Tudi v tem orodju je ponekod prihajalo do napak. Napačno so bili zaznani soglasniki (s, š, t, k ipd). Vtičnik jih je predlagal kot samostojne note. Zaradi grafičnega vmesnika orodja je bila zaznava takih primerov enostavna. Ti so bili ročno odstranjeni iz rezultatov.

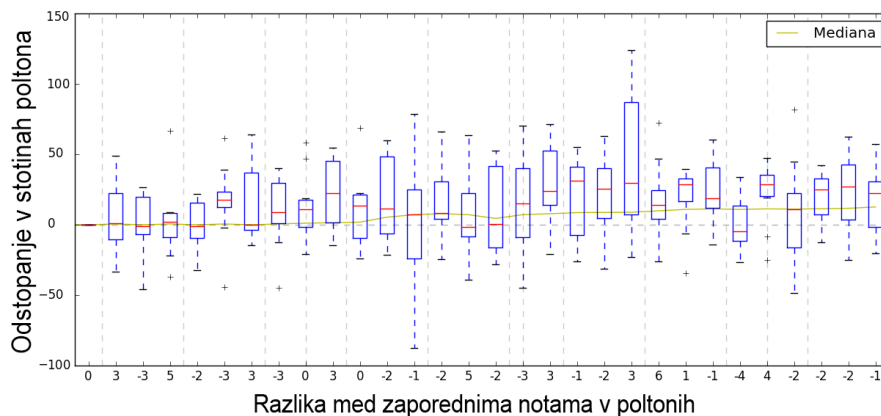
Sonic Visualiser omogoča izvoz podatkov v datoteko. Izvoženi podatki so bili analizirani s programom *PrikazRezultatov1.py*. Program prebere vse podatkovne datoteke. Pri branju jih primerja z notnim zapisom, razlike pa shrani v slovar. V zadnjem koraku te razlike izriše v grafu.

Iz grafične predstavitve je mogoče opaziti, da so bile napake pri velikih intervalih in ponavljajočih se tercah napačno predvidevane. Odstopanje pri rezultatih se sicer pojavi, vendar ni večje kot v drugih primerih. Največje napake se pojavljajo pri notah, ki pri različnih verzijah v ljudskem petju odstopajo za en polton.

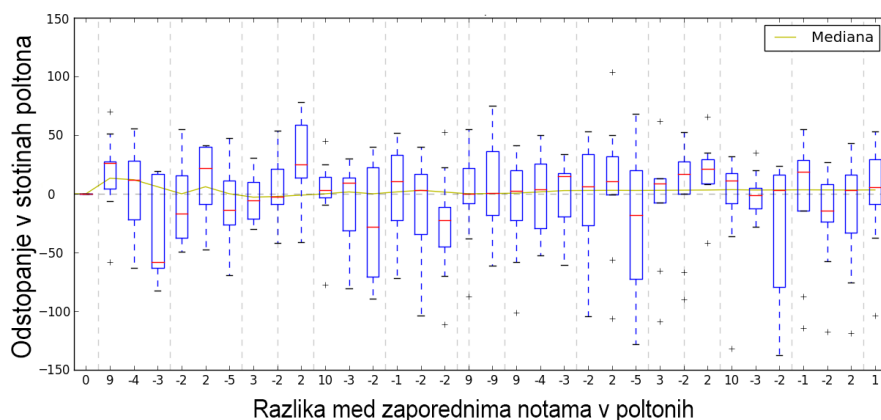
Sliki 5.3 in 5.4 prikazujeta odstopanje petja od notnega zapisa. Na vodoravni osi so za celotno melodijo zapisani intervali v poltonih. Prekinjene navpične črte sive barve predstavljajo taktne vodoravnice v notnem zapisu.

Vrsta grafa se imenuje škatla z brki (Angl. "box plot"). Poglejmo poljubno škatlo (pravokotnik). Spodnja vodoravna črta, ki je s prekinjeno navpično črto povezana s pravokotnikom označuje minimum, spodnja stranica pa prvi kvartil. Vodoravna rdeča črta predstavlja mediano (drugi kvartil). Zgornja stranica označuje tretji kvartil, vodoravna črta nad zgornjo

stranico pa maksimum. Pri nekaterih intervalih vidimo tudi znak "+". Te vrednosti se imenujejo osamelci, ker so preveč oddaljene od ostalih.



Slika 5.3: Odstopanje od notnega zapisa pri skladbi Tam dol na ravnem polju.

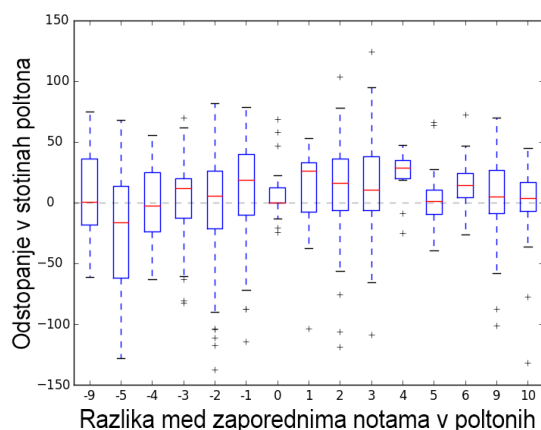


Slika 5.4: Odstopanje od notnega zapisa pri skladbi Veseli pastir.

Primer dobrega rezultata bi bil, ko je pravokotnik majhen in je razdalja med minimumom ter maksimumom čim manjša. Hkrati pa naj bi se mediana nahajala čim bližje vrednosti 0 na navpični osi. Za slab rezultat obstajata dve možnosti. Pri eni je mediana zelo oddaljena od vrednosti 0. V takem

primeru je večina pevcev zgrešila ton. Pri drugi možnosti se slab rezultat opazi kot velika razdalja med minimumom in maksimumom ter kot velik pravokotnik. Tak rezultat pomeni, da so bile razlike med pevci zelo velike.

V analizi je za referenčno noto izbran prvi ton pevca. Iz referenčne note so, s pomočjo ročno napisanega zaporedja, izračunane vrednosti za celotno skladbo. Z njimi primerjamo vrednosti odpetih not. Skupne napake, za posamezen interval pri obeh skladbah, so prikazane na sliki 5.5. Iz slike vidimo, da noben interval po višini posebno ne izstopa, razen velika terca (4 poltoni). Izstopajoči interval najdemo pri vrednosti 4 na vodoravni osi. Ta pa ni vplivala na premik intonacije, ker se pojavi le na koncu skladbe Tam dol na ravnem polju (slika 5.3). Skupna mediana se je začela spreminjati že pred tem intervalom.



Slika 5.5: Napake za posamezen interval pri obeh skladbah.

Mediana spreminjanja intonacije na koncu skladbe Veseli pastir znaša 4 stotine poltona. Nad izmerjenimi podatki smo pozitivno presenečeni. Tako majhne razlike s prostim ušesom namreč skorajda ni mogoče slišati. Pri tonu s frekvenco 440 Hz znaša ta razlika le 1.3 Hz.

Pri skladbi Tam dol na ravnem polju znaša mediana 13 stotin, kar sovпада s podatki v literaturi [1]. Tam je bila izmerjena mediana 11 stotin. Zapisan

je tudi podatek o izmerjenih vrednostih pri različnih zborih. Te znašajo od 10 do 15 stotin. Ugotovitve avtor v [1] označi kot dodatne dokaze, da so premiki v intonaciji lahko odvisni od razlike med naravno in temperirano uglasitvijo.

5.2 Iskanje napak pri razlikah med uglasitvami

V nadaljevanju je bil posnet kvartet pevcev, ki je odpel pevsko vajo (slika 5.6), kjer naj bi bili premiki v intonaciji izraziti. Pri našem notnem zapisu smo zamenjali melodiji soprana in tenorja. Avtor v [13] omeni, da je melodija za glas tenor zapisana precej visoko. S to spremembo smo želeli izključiti vpliv na intonacijo zaradi visoke lege pri petju. Ta del raziskave je namenjen iskanju premikov v intonaciji, ki bi jih lahko povzročile razlike med naravno in temperirano uglasitvijo.

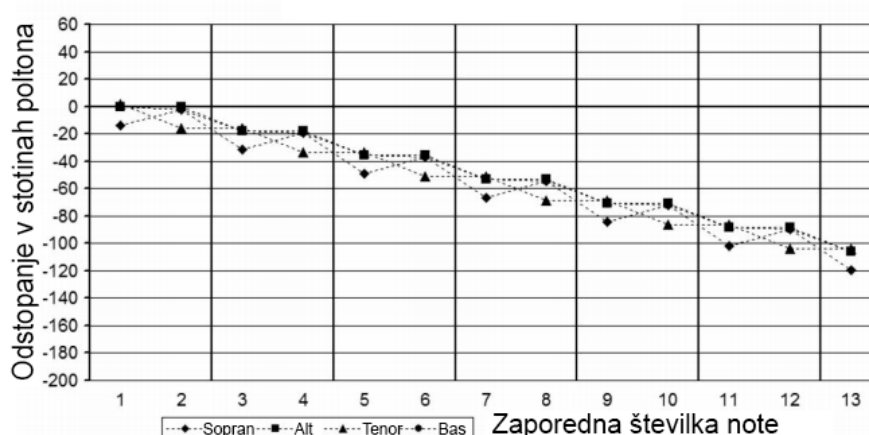


Slika 5.6: Notni zapis vaje.

V [13] je snemanje potekalo z elektrodami na glasilkah, kar je izničilo vpliv ostalih pevcev. Nam taka oprema ni bila dostopna. Pevci so bili posneti s štirimi mikrofoni. Zvok iz vsakega mikrofona je bil posnet z digitalnim snemalnikom Zoom H5, ki omogoča štiri kanalno snemanje.

Vpliv zvokov ostalih pevcev je bil zmanjšan v programu Audacity. Določili smo amplitudni prag, pod katerim so bile vrednosti odstranjene. S tem smo

zagotovili, da med posameznimi notami niso slišni ostali trije pevci. Posnetki so bili nato obdelani s programom *AnalizaPosnetkov.py*. Rezultati so bili shranjeni v tekstovno datoteko. S programom *PrikazRezultatov2.py* so bile za shranjene rezultate izračunane napake pri petju. Te napake program na koncu še izriše v grafu.

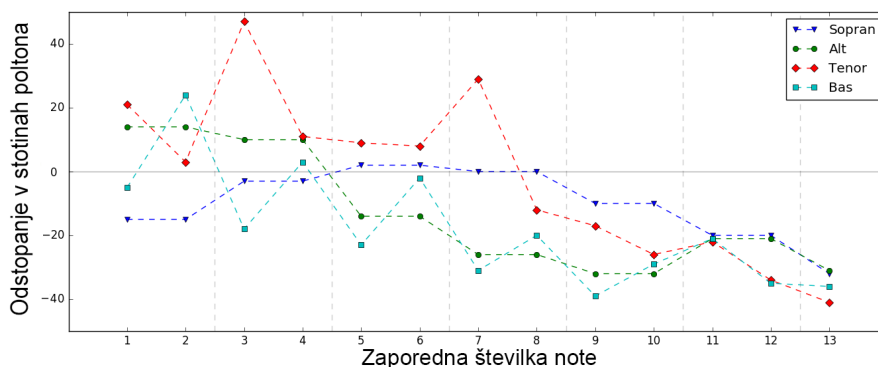


Slika 5.7: Predviden premik pri petju vaje [10].

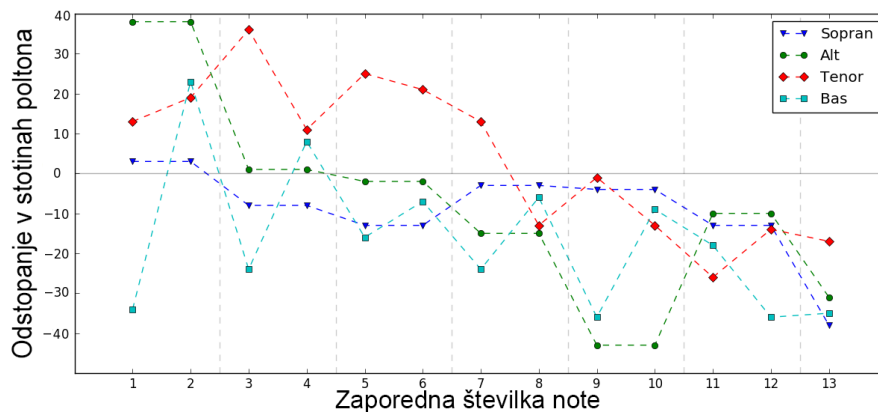
Rezultati eksperimenta so prikazani na slikah 5.8 in 5.9. Že ob primerjavi s predvidenim premikanjem intonacije pri petju (slika 5.7) opazimo, da se rezultati razlikujejo. V literaturi je praktični poizkus boljše sovpadal s predvidenimi vrednostmi [13]. V naših rezultatih je mogoče opaziti podobnosti le pri pevcu najnižjega glasu³, kjer so rahli premiki navzdol enakomerni.

Pri obeh ponovitvah je mogoče opaziti znižanje intonacije v drugi polovici vaje. Ta del še najbolj sovpada s predvidenimi premiki. Končni premik v intonaciji je izrazito manjši od predvidenega. Srednja vrednost zadnjega akorda v prvem snemanju odstopa za -37.5 , pri drugem pa za -33 stotin poltona. Oba rezultata sta bistveno manjša od predvidenih -100 stotin. Po naši oceni so rezultati tako dobri, ker se pevci zavedajo problema spreminjanja intonacije.

³Bas je najnižji (moški) glas.



Slika 5.8: Dejanski premik - prva ponovitev vaje.



Slika 5.9: Dejanski premik - druga ponovitev vaje.

V prvi polovici posnetkov se povprečje vrednosti giblje okoli ničle. Posamezne note so sicer precej visoke ali nizke. V drugi polovici se vrednosti približajo druga drugi. V eni točki so vrednosti celo zelo blizu (slika 5.8, zaporedna številka note 11). Približane vrednosti pomenijo, da so note akordov točne in akordi zvenijo prijetno. Iz takih rezultatov sklepamo, da so med vajo pevci začeli poslušati zvenenje akorda, ki so ga sestavljale njihove note. Kot je bilo omenjeno v poglavju 3, se pevci v skupini nagibajo k petju v naravni uglasitvi. Torej, pri naši analizi, v drugi polovici vaje, drsenje intonacije navzdol potrjuje ugotovitve v literaturi: pri petju se zgodijo premiki intonaciji, če pevci pojejo v naravni uglasitvi [13].

Poglavje 6

Sklepne ugotovitve

V diplomski nalogi je bil narejen pregled nad metodami za določanje frekvence signala. Podrobneje je bilo raziskano delovanje avtokorelacije in glavne značilnosti metode pYIN, ki prav tako temelji na avtokorelaciji. Glede točnosti petja obstaja kar nekaj znanstvenih člankov. Vključena je bila tista literatura, ki se je najbolj navezovala na to temo. V diplomski nalogi smo se posvetili iskanju razlogov za spreminjanje intonacije pri petju. Naš cilj je bil analizirati problem spreminjanja intonacije. S problemom se srečuje večina vokalnih sestavov brez glasbene spremljave.

6.1 Pregled praktičnega dela in rezultatov

V praktičnem delu so bili analizirani posnetki posameznih pevcev, ki so zapeli dve skladbi. Kvartet pa je istočasno odpel pevsko vajo, kjer naj bi intonacija od začetka do konca zdrsnila za pol tona navzdol.

V prvem delu praktičnega dela raziskave so bili posnetki posameznih pevcev analizirani v orodju Audacity s pYIN vtičnikom. Izvoženi podatki so bili obdelani v programskem jeziku Python s skripto *PrikazRezultatov1.py*. Obdelava je zavzemala primerjavo posnetkov z notnim zapisom, računanje odstopanja in grafični prikaz rezultatov. Za drugi del sta bili napisani dodatni dve skripti. Prva (*AnalizaPosnetkov.py*) je analizirala posnetke kvar-

teta tako, da je posnetke razdelila na posamezne okvirje in iz njih izračunala osnovne frekvence. Te so bile pretvorjene v MIDI zapis za lažjo predstavitev. Druga (*PrikazRezultatov2.py*) je, podobno kot v prvem delu, računala odstopanja od notnega zapisa in podatke izrisala. S to razliko, da je bila prilagojena za pevsko vajo. Notni zapisi med glasovi so se namreč razlikovali tako v melodiji kot v ritmu.

Prvi del je namenjen iskanju povezav med spreminjanjem intonacije pri določenih glasbenih intervalih. V splošnem smo s kvaliteto petja zadovoljni. V enem primeru se intonacija skoraj ni spremenila, v drugem pa se ne razlikuje bolj kot v literaturi. Pri posameznih intervalih, predvsem pri skladbi Veseli pastir, največje napake znašajo tudi en polton in več. Napake pripisujemo dejstvu, da nekateri pevci skladbe niso tako dobro poznali. Kljub takim napakam znaša srednja vrednost napak pri skladbi Veseli pastir le 4 stotine poltona. Napako lahko zanemarimo, saj je tako majhna, da bi jo komaj opazil tudi poznavalec glasbe. Pri ljudski skladbi Tam dol na ravnem polju znaša srednja vrednost napak 13 stotin poltona. Kar sovпада z vrednostmi iz druge literature. Iz rezultatov sklepamo, da napake pri posameznih intervalih ne predstavljajo tako velikega vpliva na intonacijo.

Pevec, ki poje točno, ni nujno definiran kot dober pevec. Poleg točnosti je pomemben tudi način interpretacije, kvaliteta glasu, način izdelovanja fraze ipd. Točnost je izmed vseh odlik pevca ta, ki jo je mogoče enostavno izmeriti. Izmerjeno petje za prvo skladbo ocenjujemo kot zelo dobro, za drugo pa kot dobro.

Zanimivo bi bilo ponoviti izračun s preoblikovanim programom, ki bi hranil rezultate vsakega pevca posebej. Z izračunom srednjih vrednosti bi dobili oceno natančnosti petja posameznega pevca. To bi nam omogočalo primerjavo not najbolj in najmanj točnega pevca. Primerjava bi pokazala, če se napake pojavljajo na istih mestih ter kakšne so razlike med različno nadarjenimi pevci. S tem primerom bi bil analiziran še eden vzrok spreminjanja intonacije - izkušnost pevca.

Drugi del je namenjen analizi napak pri večglasnem petju. Tukaj bi na

intonacijo lahko vplivala razlika med naravno in temperirano uglasitvijo. V literaturi smo zasledili notni zapis pevske vaje, ki je sestavljena iz postopnih kromatičnih prehodov. Prvi in zadnji akord sestavljajo isti toni, ki so eno oktavo narazen. Zaradi razlik med uglasitvami je predviden premik intonacije za pol tona navzdol (-100 stotin). To je bilo v literaturi tudi potrjeno s praktičnim preizkusom. V tej nalogi je bil izmerjen premik -37.50 stotin pri prvem in -33 stotin pri drugem snemanju. Rezultata kažeta na to, da so pevci v našem primeru peli veliko bolje. Po koncu vsakega snemanja so pevci sami ocenili potek petja. V obeh primerih so komaj opazili premike v intonaciji. To je dodaten dokaz, da so izmerjene razlike majhne.

Opazili smo podobnosti pri premiku intonacije od polovice notnega zapisa do konca. V prvi polovici podatki niso primerljivi s predvidenimi. Vrednosti so bile med seboj oddaljene, pojavljali so se tudi premiki navzgor. Nastanek teh premikov je mogoče razložiti s tem, da se pevci zavedajo premikov navzdol in jih skušajo preprečiti s čim višjim petjem. Po določenem času se vendarle vrednosti približajo druga drugi. Takrat so pevci začeli poslušati zvenenje akordov. Peli so tako, da so bili akordi uglaseni v naravni intonaciji, takrat zvenijo najbolj prijetno. V tistem trenutku se je intonacija začela premikati navzdol, kot je bilo predvideno. To dokazuje, da je razlika med naravno in temperirano uglasitvijo tehten razlog za premike v intonaciji. Rezultati potrjujejo tudi dejstvo, da pevec strmi k petju v naravni uglasitvi.

6.2 Analiza rezultatov in izboljšave

V veselje nam je, da smo imeli priložnost analizirati prav pevce v zboru, kjer se je pojavila motivacija za naše delo. Glede izmerjenih rezultatov petja lahko zaključimo, da so v prvem delu rezultati pri eni skladbi sovpadali s predvidenimi vrednostmi iz literature. Pri drugi je bilo petje celo bolj točno in bližje notnemu zapisu. V drugem delu so bili rezultati prav tako boljši od predvidenih, saj je bil premik med prvim in zadnjim akordom za dve tretjini manjši. Natančnejši rezultati bi bili v prvem delu mogoči z analizami daljših

skladb ali ponovitvami skladbe v istem posnetku. Pri večglasnem petju so bili vsi štirje sodelujoči solopevci (od nekaj mesecev do nekaj let solopetja). Vredno bi bilo analizirati tudi različne kombinacije pevcev z različnimi nivoji pevskih izkušenj.

V izmerjenih rezultatih smo opazili določene napake, vendar so v vseh primerih rezultati boljši od pričakovanih. Večje napake so bile opažene pri skladbi Veseli pastir, ki je pevcem manj znana in po sestavi zahtevnejša. Iz tega smo ugotovili, da sta pomembna dejavnika pri točnosti petja tudi poznavanje skladbe in samozavest pri petju.

6.3 Praktičen pomen ugotovitev

Pomembna ugotovitev v nalogi je, da za spreminjanje intonacije ni kriva izključno nenadarjenost pevca. Nadarjenost je le eden od dejavnikov. V nalogi smo potrdili izjemen vpliv na intonacijo, ki se kaže v razliki med naravno in temperirano uglasitvijo. V rezultatih drugega dela analize vidimo zanimiv pojav. Na polovici vaje pevci prisluhnejo drug drugemu. Izmerjene vrednosti se približajo, note so bolj točne, akordi pa čisto uglašeni. Takrat se začne intonacija premikati v predvideni smeri - navzdol. To potrjuje, da se intonacija v skladbi spreminja zaradi razlik med naravno in temperirano uglasitvijo.

Iz te ugotovitve se lahko učijo tudi ostali zbori. K kakovostnejšemu poustvarjanju glasbe lahko pripomorejo tako, da razloge za spreminjanje intonacije ne iščejo pri pevcih, temveč v skladbi. V podpoglavju 2.2.3 se nahajajo tudi nekateri uporabni nasveti, ki pripomorejo k doseganju bolj konsonantnega zvoka v zboru.

Literatura

- [1] Matthias Mauch, Klaus Frieler, and Simon Dixon. Intonation in unaccompanied singing: Accuracy, drift and a model of reference pitch memory. *Journal of the Acoustical Society of America*, 2014.
- [2] Philip Mcleod, Damon Simpson, Robert Visser, Mike Phillips, Ignas Kukenys, Yaoyao Wang, Arthur Melissen, Natalie Zhao, and All The. Fast, accurate pitch detection tools for music analysis, 2008.
- [3] Habib Hajimolahoseini, Rassoul Amirfattahi, Saeed Gazor, and Hamid Soltanian-Zadeh. Robust estimation and tracking of pitch period using an efficient Bayesian filter. 24(7):1219–1229, July 2016.
- [4] Razlika med časovno in frekvenčno domeno. [Online]. Dosegljivo: <http://engineersknowledge.blogspot.si/2013/03/sine-wave-in-frequency-domain.html>. [Dostopano 23. 1. 2017].
- [5] Onur Babacan, Thomas Drugman, Nicolas d’Alessandro, Nathalie Henrich, and Thierry Dutoit. A comparative study of pitch extraction algorithms on a large variety of singing sounds. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 7815–7819. IEEE, 2013.
- [6] Vamp vtičniki. [Online]. Dosegljivo: <http://www.vamp-plugins.org/>. [Dostopano 20. 1. 2017].
- [7] C. P. Singh and T. K. Kumar. Efficient pitch detection algorithms for pitched musical instrument sounds: A comparative performance

- evaluation. In *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 1876–1880, Sept 2014.
- [8] France Mihelič. Avtokorelacija naključnih signalov. [Online]. Dosegljivo: <http://luks.fe.uni-lj.si/sl/studij/AS/pub/AS8.pdf>. [Dostopano 20. 1. 2017].
- [9] Matthias Mauch and Simon Dixon. PYIN: A fundamental frequency estimator using probabilistic threshold distributions. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 659–663. IEEE, 2014.
- [10] David M Howard. Equal or non-equal temperament in a capella satb singing. *Logopedics phoniatriks vocology*, 32(2):87–94, 2007.
- [11] Richard Seaton, David Sharp, and Dennis Pim. Pitch drift in a capella choral singing: the outcomes from an international survey. *Proceedings of Institute of Acoustics*, 36(3):312–319, 2014.
- [12] Hiroko Terasawa. Pitch drift in choral music. 2004.
- [13] Howard David M. Intonation drift in a capella soprano, alto, tenor, bass quartet singing with key modulation. *Journal of Voice*, 21(3):300–15, 2007.
- [14] Sovpadanje harmonikov. [Online]. Dosegljivo: <https://needhamia.com/tunesing/index.htm>. [Dostopano 21. 1. 2017].
- [15] Kvintni krog. [Online]. Dosegljivo: <http://www.music-theory-for-musicians.com/circle-of-fifths.html>. [Dostopano 21. 1. 2017].
- [16] Lestvice, naravna in temperirana uglasitev. [Online]. Dosegljivo: <http://www.phy.mtu.edu/~suits/scales.html>. [Dostopano 21. 1. 2017].

-
- [17] David Wright. *Mathematics and music*, volume 28. American Mathematical Soc., 2009.
- [18] Wikipedija: Frekvence na klavirju. [Online]. Dosegljivo: https://en.wikipedia.org/wiki/Piano_key_frequencies. [Dostopano 21. 1. 2017].
- [19] Audacity. [Online]. Dosegljivo: <http://www.audacityteam.org/about/features/>. [Dostopano 26. 1. 2017].
- [20] Sonic visualiser. [Online]. Dosegljivo: <http://www.sonicvisualiser.org/>. [Dostopano 26. 1. 2017].
- [21] Programiranje v pythonu. [Online]. Dosegljivo: <https://www.python.org/>. [Dostopano 26. 1. 2017].
- [22] Audiolab paket. [Online]. Dosegljivo: <https://pypi.python.org/pypi/scikits.audiolab/>. [Dostopano 26. 1. 2017].
- [23] Numpy paket. [Online]. Dosegljivo: <http://www.numpy.org/>. [Dostopano 26. 1. 2017].
- [24] Scipy paket. [Online]. Dosegljivo: <https://www.scipy.org/index.html>. [Dostopano 26. 1. 2017].
- [25] Scipy paket. [Online]. Dosegljivo: <https://www.scipy.org/about.html>. [Dostopano 26. 1. 2017].
- [26] Matplotlib knjižnica. [Online]. Dosegljivo: <http://matplotlib.org/>. [Dostopano 26. 1. 2017].
- [27] O programski opremi audacity. [Online]. Dosegljivo: <https://www.scipy.org/about.html>. [Dostopano 26. 1. 2017].
- [28] Beauregard Gerry. Flash aplikacija za detekcijo višine tona. [Online]. Dosegljivo: <https://gerrybeauregard.wordpress.com/2013/04/30/pitch-detection-in-flash/>. [Dostopano 20. 1. 2017].